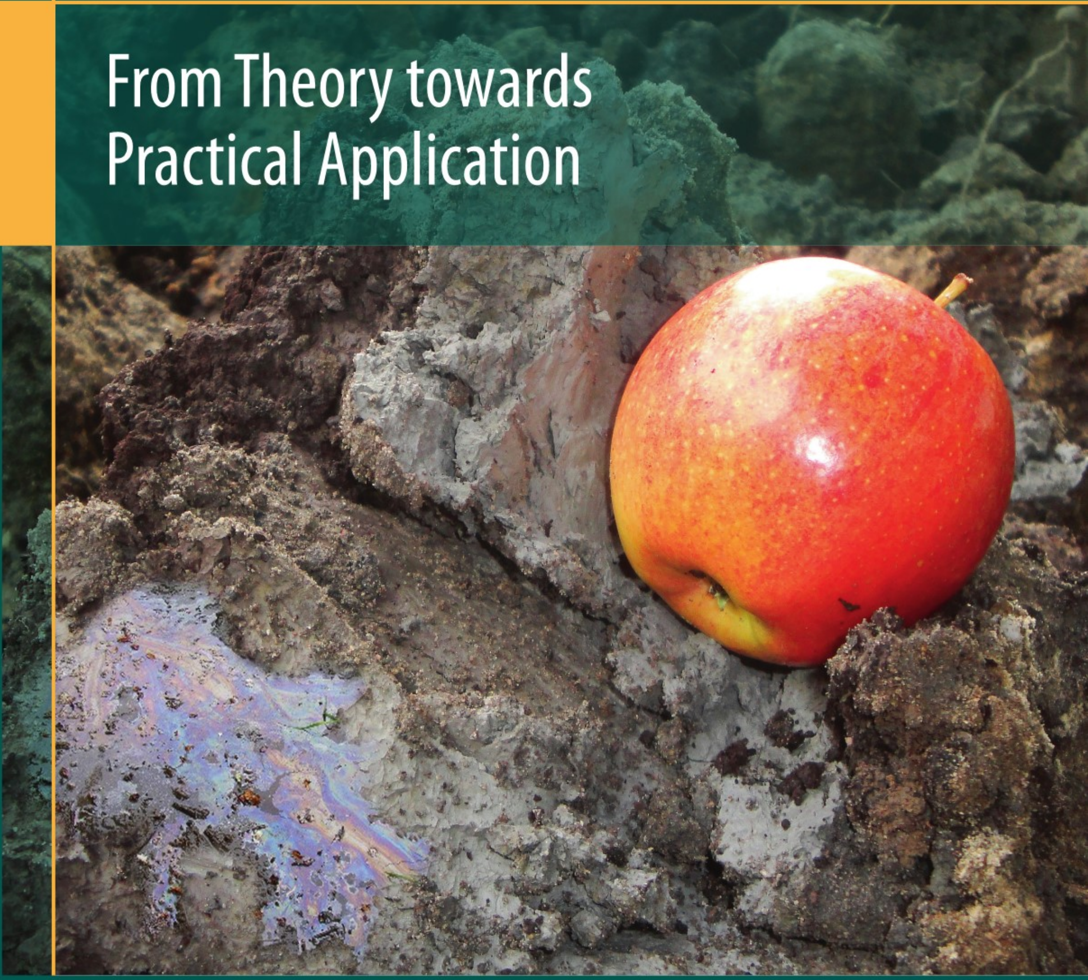
Frank A. Swartjes *Editor*

Dealing with Contaminated Sites

From Theory towards Practical Application



Springer

Dealing with Contaminated Sites

About the Editor

Frank A. Swartjes PhD (1960) has been working in the field of risk-based contaminated site management since 1986. In this year, he graduated in Soil Science and Hydrology from Wageningen University, The Netherlands, and, in 1990, received his PhD on The Fate and Transport of Contaminants in Soil and Groundwater (*cum laude*) from the Technical University of the former West-Berlin, Germany. Subsequently, he spent two interesting years with the consultancy firm of Fugro-McClelland Engineers in The Netherlands, where he worked on contaminated site investigations and remediations. In 1991, he joined the National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands, where he is still employed today. At RIVM he has contributed to the foundation of the Dutch Soil Protection Act (related to Human Health, Ecological and Groundwater-related Risk Assessment), both as a project manager and as a scientist. Moreover, he has been involved in the harmonisation of Risk Assessment tools for soil contamination in the European Union and in knowledge transfer to countries that are in the early stages of risk-based contaminated site management. His general interest is in implementing in-depth scientific knowledge for practical procedures and cost-effective solutions. He is a frequently invited speaker at conferences.

Dealing with Contaminated Sites

From Theory Towards Practical Application

Edited by

Frank A. Swartjes

National Institute of Public Health and the Environment (RIVM),
Bilthoven, The Netherlands

Springer

Editor
Frank A. Swartjes
National Institute of Public Health
and the Environment (RIVM)
P.O. Box 1
3720 BA Bilthoven
The Netherlands
frank.swartjes@rivm.nl

ISBN 978-90-481-9756-9 e-ISBN 978-90-481-9757-6
DOI 10.1007/978-90-481-9757-6
Springer Dordrecht Heidelberg London New York

Cover illustration: Soil material and an apple at a contaminated site in Utrecht, the Netherlands; photo by F. Swartjes (artistic support by M. Kerkhoven, Ch. Langenberg of LangVorm and A. Verschoor)

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Dealing with contaminated sites:
When you have a wide scientific knowledge concerning contaminant fate and transport processes in soil and groundwater, site characterization, Human Health Risk Assessment, Ecological Risk Assessment and Groundwater-related Risk Assessment, AND have experience with designing cost-efficient Risk Management solutions AND have a creative personality AND have good communication qualities AND are in a position to take policy decisions: Go ahead.
Otherwise: Build a team.

Frank Swartjes

Preface

In 2006, two paths that led to this book came together at a conference in San Diego, California. First, the vague idea that I had for some years about preparing this book began to evolve into a mature plan. Second, a representative of Springer publishers contacted me, asking whether I would be interested in preparing a book on contaminated site management.

My plans for this book on contaminated site management were simply a response to the many questions on the subject that I have received, and continue to receive, almost on a daily basis. The senders of these questions can roughly be subdivided into two categories. The first category of e-mailers consists of individuals who are not very familiar with contaminated site management, but are in the early stage of risk-based contaminated site management or have plans for developing a contaminated site management framework. These individuals are often from the newer or candidate countries of the European Union, or from Asia, Latin America and Africa. The second category of people that contact me regularly is made up of practitioners who know about contaminated site management, often even having a considerable amount of knowledge about it, but who have specific detailed questions. My conclusion then was that it would be useful to bundle all contaminated site management information, theory and possibilities for practical application, for both these target groups in one handbook.

This plan was further nourished by the backup I have received from individuals in my international network, both technically as well as personally, over the past years. Among the colleagues from the European networks CARACAS, CLARINET, HERACLES and NICOLE, and colleagues from the USA, Canada and Australia, some kind of brotherhood has developed. It has been a continuous process in which the procedures for contaminated site management have undergone a major evolution that has resulted in the sound and efficient procedures we use today. None of us would have been individually able to do this without the scientific and social stimulus of our colleagues. This book is a common product, indirectly the fruit of our collaboration over the last two decades, since the majority of the authors originate from these networks and congresses and workshops in the USA and Australia. However, a former colleague from the second half of the 1980s – and my supervisor when I was a young trainee in Riverside, California, in 1984 – is also among the contributing authors. I am happy and feel privileged to have all the contributions of these individuals in this one book; the author list includes many of the world's most acknowledged experts in their specific fields, some of whom I consider to be personal friends. Last but not least, I feel the need to mention the late Prof. Colin Fergusson, PhD, one of the pioneers with regard to contaminated site management, along with the sweet and talented late Dr. Arantzazu Ursulai, and I wish to pay some kind of tribute to them. Their contributions should have been in this book, had their lives not come to an end far too prematurely.

I sincerely hope this book will contribute to a sounder and more efficient way of dealing with contaminated sites, all over the world.

Utrecht, The Netherlands Frank Swartjes

Acknowledgments

Putting the book together has been an intriguing, but time-consuming experience. I owe a debt of gratitude to a few people in my direct and indirect environment. First of all, Mr. Ruud Cino, director of the Soil Remediation Department of the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM). He believed in the book project and in me and, moreover, proved to be a source of inspiration over the last years in several interesting projects. Additionally, I want to thank the late Dr. Peter Doelman, who supported the idea of this book with his characteristic style and enthusiasm, both before and during the book process. It is hard to realise, but ironically Peter passed away a few weeks before the book was ready. I am also grateful to Prof. Dr. Paul Bardos, manager of r^3 Environmental Technology Ltd in the UK, Dr. Christopher Teaf of the Florida State University in the USA, and Prof. Dr. Bob Harris of the University of Sheffield in the UK, for their technical, linguistic or moral support during the book project. Besides, I would like to thank my superiors at the National Institute of Public Health and the Environment (RIVM); these are Prof. Dr. Ton Breure (head of the Laboratory of Ecological Risk Assessment), and Dr. Reinout Wottiez (director of the Environment and Safety Division). I thank them for temporarily freeing up my contract with RIVM and, hence, offering me a flexible time schedule in order to work on this book dealing with contaminated sites. And not least, I would like to sincerely thank all authors who contributed to this book. I realise very well that quite an effort was required from these experts, who generally belong to the busiest individuals in our field. Their dedication and enthusiasm was a blessing and I am proud of the result – this book, Dealing with Contaminated Sites, that we delivered together.

Mieke was enthusiastic about my book idea from the beginning to the very end. She gladly paid the price by letting me spend all those hours on this book project and, hence, provided invaluable moral support. I knew it already, but this has proved beyond a shadow of a doubt that she is the most tolerant person in this world.

Utrecht, The Netherlands Frank Swartjes

Reviewers

It has been a major asset that many contaminated site management experts, be it befriended colleagues or people I have never met before, have shown themselves prepared to search deeply into their colleagues' texts. These reviewers have made numerous, useful suggestions, ranging from simple typos on up to whole philosophical explanations. In some cases, this has resulted in intensive e-mail discussions between reviewers, the authors and myself. Many draft Chapters benefited immensely from these comments. Therefore, I would like to express my deepest appreciation to those experts who provided useful comments to the draft chapters:

Dr. Jack Adams (University of Utah, Salt Lake City, UT, USA), Dr. Sebastian Bauer (Christian-Albrechts-University, Kiel, Germany), Mr. Andreas Bieber (Ministry of Environment, Bonn, Germany), Dr. Christof Beyer (Christian-Albrechts-University, Kiel, Germany), Dr. Karen Bradham (U.S. Environmental Protection Agency, Research Triangle Park, NC, USA), Dr. Cynthia Brooks (Greenfield Environmental Trust Group, Inc. (GETG), Cambridge, MA, USA), Dr. Fred Coulon (Cranfield University, Cranfield, UK), Dr. Michelle Crimi (Clarkson University, Potsdam, NY, USA), Prof. Dr. Bill Doucette (Utah State University, Logan, UT, USA), Mr. Victor Dries (Flemish Public Waste Agency (OVAM), Mechelen, Belgium), Prof. Dr. Valery E. Forbes (Roskilde University, Roskilde, Denmark), Dr. Natalie Freeman (University of Florida, Gainesville, FL, USA), Prof. Dr. Bob Harris (University of Sheffield, Sheffield, UK), Dr. Ian Hers (Golder Associates Ltd., Burnaby, BC, Canada), Prof. Dr. Ivan Holoubek (Masaryk University, Brno, Czech Republic), Prof. Dr. Paul Johnson (Arizona State University, Tempe, AZ, USA), Prof. Dr. Andreas Paul Loibner (University of Natural Resources and Applied Life Sciences, Vienna, Austria), Dr. Richard O. Gilbert (Rockville, MD, USA), Dr. Detlef Grimski (Federal Environmental Agency (UBA), Dessau, Germany), Prof. Dr. Antonio Marcomini (University Ca' Foscari of Venice, Venice, Italy), Prof. Dr. Bernd Marschner (Ruhr-University Bochum, Bochum, Germany), Prof. Dr. Michael Matthies (University of Osnabrück, Osnabrück, Germany), Mr. Andreas Paetz (DIN German Institute for Standardization, Berlin, Germany), Ms. Anna-Maija Pajukallio (Ministry of the Environment, Helsinki, Finland), Prof. Dr. Simon Pollard (Cranfield University, Cranfield, UK), Dr. José Paulo Sousa (University of Coimbra, Coimbra, Portugal), Dr. Henning Prommer (CSIRO Land & Water, Urban and Industrial Water, Perth, Australia), Prof. Dr. R. Neil Reese (South Dakota State University, Brookings, SD, USA), Dr. Rob Reid (The University of Adelaide, Adelaide, Australia), Prof. Dr. Ken Reimer (The Royal Military College of Canada, Kingston, TO, Canada), Mr. Peter Rostron (University of Sussex, Brighton, UK), Dr. Michael V. Ruby (Integral Consulting Inc., Denver, CO, USA), Dr. Michiel Rutgers (National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands), Mr. Ebel Smidt (SG Consultancy and Mediation Ltd., Engelen, The Netherlands), Dr. Jonathan Smith (Shell, Chester, UK), Mr. Michael Smith (M.A. Smith Environmental Consultancy, Hemel Hempstead, UK), Dr. Klaus Schneider (Research and Advisory Institute for Hazardous Substances, Freiburg,

Germany), Dr. Ed Stanek (University of Massachussets, Amherst, MA, USA), Dr. Curtis Stanley (Shell Global Solutions, Houston, TX, USA), Prof. Dr. Nico M. van Straalen (Free University, Amsterdam, The Netherlands), Dr. Christopher Teaf (Florida State University, Tallahassee, FL, USA), Mr. Johan Van Veen (Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, The Netherlands), and Dr. Harold I. Zeliger (West Charlton, NY, USA).

Utrecht, The Netherlands Frank Swartjes

Contents

Part I Introduction

Part II Site Investigation

Part III Human Health Aspects

Part IV Ecological Aspects

Part V Groundwater-Related Aspects

Part VI Risk Management

Part VII Frameworks

Contributors

Markus Ackermann DuPont de Nemours, Geneva, Switzerland, markus.ackermann@che.dupont.com

Markus Ackermann PhD in hydrogeology has been working in management of water resources and contaminated land since 1991. He has been with DuPont's Corporate Remediation Group since 2005 and manages the company's portfolio of contaminated land in Europe.

R. Paul Bardos Environmental Technology Ltd, Reading, UK, paul@r3environmental.co.uk

Prof. Dr. R. Paul Bardos manages r^3 Environmental Technology Ltd (www.r3environmental.com) and is Special Professor at two UK universities. He was a participant of CLARINET, the first international project to formulate concepts of sustainable remediation. He is a member of the UK Sustainable Remediation Forum Steering Group and the NICOLE Working Group on Sustainable Remediation.

Laurent M.M. Bakker Tauw bv., Deventer, The Netherlands, laurent.bakker@tauw.nl

Laurent M.M. Bakker MSc has a background in hydrogeology and soil physics and is working on contaminated sites. He is with the Tauw Group, an independent international consulting and engineering company, the Netherlands since 1990. His expertise is in decision making and communication regarding the remediation of industrially contaminated sites.

Michel Beaulieu Quebec Ministry of Sustainable Development, Environment and Parks, Quebec, Canada, michel.beaulieu@mdeep.gouv.qc.ca

Michel Beaulieu PhD has a background in biology and has been working since 1984 on the rehabilitation of contaminated sites at the Quebec Ministry of Sustainable Development, Environment and Parks. He was involved in the elaboration of various regulations, policies and guidelines on the assessment and revitalisation of contaminated sites in Canada.

Johan Bierkens Flemish Institute of Technological Research (VITO), Mol, Belgium, Johan.bierkens@vito.be

Johan Bierkens PhD was trained as a biologist and has been involved in several programs on risk and impact assessment and the calculation of environmental standards for soil. Currently he holds a position as a research scientist at the Flemish Institute for Technological Research (VITO), where his interests focus on human exposure assessment and children's exposure.

Roseline Bonnard National Institute of Industrial Environment and Risks (INERIS), Verneuil-en-Halatte, France, Roseline.BONNARD@ineris.fr

Dr. Roseline Bonnard MEng has a background in biological and sanitary engineering and is working on human health risk assessment. She is with the National Institute of Industrial Environment and Risks (INERIS), FRANCE, since 1995. Her expertise is exposure modelling and uncertainties assessment.

Anton M. Breure National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands, ton.breure@rivm.nl

Prof. Dr. Anton M. Breure studied chemistry and microbiology. In 1986 he joined the National Institute for Public Health and the Environment (RIVM), where he is head of the Laboratory for Ecological Risk Assessment. He holds an extraordinary chair in ecological risk assessment at Radboud University in Nijmegen, the Netherlands.

Dick J. Brus Wageningen University Research Centre, Alterra, Wageningen, The Netherlands, dick.brus@wur.nl

Dick J. Brus PhD works at Wageningen University and Research Centre. His main fields of interest are sampling theory and geostatistics. He has rich research experiences in developing and applying statistical methods for natural resources inventory and monitoring. He is co-author of the book 'Sampling for Natural Resource Monitoring', which is widely acclaimed in soil, earth, environmental, agriculture, and statistical sciences.

Lucia Buve Umnicore, Brussels, Belgium, Lucia.Buve@umicore.com

Lucia Buve MSc has a background in exploration geology. She started her carrier as exploration geochemist in Africa (Gabon and Mali), Mexico and Argentina. Since more than 15 years, she is working in the field of contaminated sites. Today, she is Pollution Prevention Manager at the corporate EHS department of Belgium based Umicore, where she mainly deals with technical, legal and financial aspects of the company's contaminated sites worldwide.

Mark R. Cave British Geological Survey, Keyworth, Nottingham, UK, mrca@bgs.ac.uk

Mark R. Cave BSc, ARCS, PhD, MSc, MRSC, CSci, CChem is the Geochemical Baselines and Medical Geology Team Leader for the Environmental Geoscience Baselines programme of the British Geological Survey. As chairman of the Bioaccessibility Research Group of Europe he has research interests in validated protocols for in-vitro bioaccessibility testing.

Christa Cornelis Flemish Institute for Technological Research (VITO), Mol, Belgium, christa.cornelis@vito.be

Ir. Christa Cornelis graduated as an engineer in chemistry and agricultural industries. She works at the Flemish Institute for Technological Research (VITO) and is involved in human health risk assessment of contaminated sites since 1991. She has supported the Flemish Government since the start of its soil legislation. At present, her expertise covers the area of human exposure and risk assessment of environmental and food contaminants.

Helen E. Dawson US EPA Region 8, Denver, CO, USA, dawson.helen@epa.gov

Helen E. Dawson PhD has a background in geology, geochemistry, and contaminant transport. She has been with the US EPA since 1999, currently manages a group of scientists addressing science policy issues for the agency, and has been a key member of EPA's vapour intrusion workgroup developing guidance for the agency.

Johan de Fraye Honeywell International Inc., London, UK, johan.defraye@honeywell.com

Johan de Fraye MSc has a background in geology, and is regional director of Environmental Services Europe at CH2M HILL in the UK. He has been responsible for managing remediation projects at contaminated sites all over the world.

Sébastien Denys National Institute of Industrial Environment and Risks (INERIS), Verneuil-en-Halatte, France, sebastien.denys@ineris.fr

Dr. Sébastien Denys has a background in the soil transfer of pollutants. He has been at INERIS (French Institute for Industrial Environment and Risks) since 2002 where he developed, in collaboration with other research institutes, in vivo tests to estimate bioavailability of soil pollutants and to validate in vitro measurement of bioaccessibility.

Fien Degryse Catholic University Leuven, Heverlee, Belgium, fien.degryse@ees.kuleuven.be

Dr. Fien Degryse is a post-doctoral researcher at the Catholic University Leuven (Belgium). She teaches Groundwater Chemistry and Soil Chemistry, and serves as an Associate editor of the Journal of Environmental Quality. Her research activities focus on speciation, mobility and bioavailability of metals in terrestrial and aquatic systems.

Mark Elert Kemakta Konsult, Stockholm, Sweden, mark@kemakta.se

Mark Elert BSc has for the last 30 years worked at Kemakta Consultants Co, Sweden, with risk assessment. He has a background in mathematics and physics and is a specialist in modelling. He has been one of the key persons in developing the model for the Swedish Soil Quality Standards.

Cyril Feidt University of Nancy, Nancy, France, cyril.feidt@ensaia.inpl-nancy.fr

Prof. Dr. Cyril Feidt is a professor at the University of Nancy where he has been since 1997. He has developed strategies to assess bioavailability and transfer of micro pollutants, either organics or trace elements, to animal products. He is a member of the Physical and Chemical Contaminants and Residues panel for the French Food Safety Agency.

Peter Grathwohl University of Tübingen, Tübingen, Germany, grathwohl@uni-tuebingen.de

Prof. Dr. Peter Grathwohl has a background in hydrogeology and has been working on soil and groundwater remediation/risk assessment since 1985. He is professor for hydrogeochemistry at the Center of Applied Geosciences, Tübingen University, Germany, since 1996. His expertise is on fate and transport of organic pollutants in the subsurface environment.

Juan Grima Spanish Geological Survey (IGME), Valencia, Spain, j.grima@igme.es

Juan Grima MSc has a background in hydrogeology and has been working on contaminated sites. He has held a position in the Geological Survey of Spain (IGME) since 1988, and is an advisor to the Spanish Environment Ministry. His expertise includes environmental assessment and remediation of contaminated soil and groundwater.

Tim J.T.C. Grotenhuis Wageningen University, Wageningen, The Netherlands, tim.grotenhuis@wur.nl

Tim J.T.C. Grotenhuis PhD has a background in chemical technology and microbiology and has been working in soil and sediment remediation since 1991. He is with the Section of Environmental Technology of Wageningen University, the Netherlands, since 1991. His expertise is in developing new remediation technologies and contaminant availability issues.

Ian Heasman Taylor Woodrow Developments Ltd., Newbury, UK, ian.heasman@taylorwimpey.com

Ian Heasman has an earth sciences background and has experience in industry, consultancy and academia within the area of environmental assessment and management, specialising in land quality, since 1990. His main area is brownfield regeneration.

Thomas Held Arcadis GmbH, Darmstadt, Germany, t.held@arcadis.de

Thomas Held PhD is a microbiologist and has 17 years of experience in environmental engineering. Key qualifications: MNA, R&D, advanced remediation technologies; in-situ-technologies. He was involved in several R&P projects of the KORA program. He is member of several national and international professional bodies. Most of his work has been published in numerous papers and oral conference presentations.

Mark E. Hodson University of Reading, Reading, UK, m.e.hodson@reading.ac.uk

Prof. Mark E. Hodson PhD is head of the Department of Soil Science, University of Reading, the UK's only dedicated soil science department. His expertise covers low temperature biogeochemistry, particularly relating to metals at contaminated sites. His current research includes in-situ remediation and the ecology of former mine and Brownfield sites.

Norbert Hüsers Institute of Waste Management and Contaminated Site Treatment, Technische Universität Dresden, Pratzschwitzer Str. 15, 01796 Pirna, Germany, norbert.huesers@tu-dresden.de

Dr. Norbert Hüsers After finishing his PhD in the field of analytical chemistry Norbert Hüsers worked several years in an engineering company. There he was responsible for contaminated site investigation, risk assessment and site remediation. In 2001 he was appointed to a position as scientist at the Technical University Dresden. Since that time his research work has focused on in-situ processes of organic contaminants, especially on the biodegradation of these substances.

Diederik Jacques Belgian Nuclear Research Centre SCK•CEN, Mol, Belgium, djacques@sckcen.be

Dr. Diederik Jacques is a researcher at the Belgian Nuclear Research Centre SCK•CEN. His expertise is in experimental and numerical studies on water flow, solute transport and geochemical modelling in variably-saturated porous media, including soils and engineered barriers for radwaste disposal. He has cooperated in the development of the HP1 model, a powerful tool for reactive transport modelling.

Roger Jacquet Solvay, Brussels, Belgium, roger.jacquet@solvay.com

Roger Jacquet environmental expert, has 15 years experience with industrial contaminates site management projects for the international chemical Solvay Group. His focus has been on in situ management for risk control.

John Jensen University of Aarhus, Silkeborg, Denmark, jje@dmu.dk

John Jensen PhD has a background in terrestrial ecotoxicology and ecology and has been involved in various aspects of ecological risk assessment of contaminated sites for more than 15 years. He has been at the National Environmental Research Institute, Aarhus University, Denmark, since 1994.

Celia Jones Kemakta Konsult, Stockholm, Sweden, celia@kemakta.se

Celia Jones PhD works at Kemakta Consultants Co, Sweden. She has over 25 years experience with the assessment of health risks and environmental effects in contaminated terrestrial and aquatic environments. Her work has included risk assessments for a number of remediation projects and the development of risk assessment methods.

Catherine Jondreville University of Nancy, Nancy, France, catherine.jondreville@ensaia.inpl-nancy.fr

Catherine Jondreville has since 2000 been working as a pig and poultry nutritionist at the French National Institute for Agricultural Research. She specialises in trace elements nutrition and has developed strategies to assess bioavailability of these compounds in vivo. Since 2006, she has enlarged her activities to the transfer of micro pollutants to animal products.

Harald Kasamas Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLFUW), Vienna, Austria, harald.kasamas@lebensministerium.at

Harald Kasamas MSc has been working in contaminated land management with the Austrian Environment Ministry since 1990. He is involved in various international and European projects and networks with emphasis on contaminated land strategies. From 1996 till 2001, he was project manager of the EU Concerted Actions CARACAS and CLARINET.

Frank P.J. Lamé Deltares, Utrecht, The Netherlands, frank.lame@deltares.nl

Frank P.J. Lamé MSc has been working in the field of soil pollution, investigation, management and prevention of contamination since 1985. His technical expertise lies in the development of investigation and assessment strategies. He is active in the field of standardization, amongst others being the Chairman of ISO/TC 190 ‘Soil Quality’ as well as CEN/TC 345 ‘Characterization of Soils’.

Andrew Langley Population Health Queensland, Maroochydore, QLD, Australia, Andrew_Langley@health.qld.gov.au

Dr Andrew Langley MBBS, Grad Dip Health Ed, FAFOM, FAFPHM has a background in public health medicine and occupational and environmental medicine and works in Queensland, Australia. He has worked extensively on the development of Australian national policy for the assessment and management of contaminated land and was the principal author of enHealth’s national ‘Environmental Health Risk Assessment’ guidelines.

Charlotte N. Legind Department of Environmental Engineering, Technical University of Denmark, Lyngby, Copenhagen, Denmark, chanl@env.dtu.dk

Dr. Charlotte N. Legind made her PhD in environmental chemistry at the University of Copenhagen and is now employed as Post Doc at the Technical

University of Denmark. Her research areas are risk assessment, uptake of organic chemicals in plants, model application and development.

Ulrich Maier University of Tübingen, Tübingen, Germany, uli.maier@uni-tuebingen.de

Dr. Ulrich Maier has a background in hydrogeology and soil science and is working on computer modelling of natural attenuation of contaminants in soil and groundwater, as well as hydrogeochemical processes in general, since 1998. He is with the Center for Applied Geoscience at the University of Tübingen, Germany.

Dirk Mallants Belgian Nuclear Research Centre SCK•CEN, Mol, Belgium, dmallant@sckcen.be

Dr. Dirk Mallants is a senior researcher at the Belgian Nuclear Research Centre SCK•CEN. He has almost twenty years experience in characterizing and modelling water flow and contaminant transport in complex environments, including unsaturated soils and aquifers. At SCK•CEN he leads new developments in biogeochemical transport modelling in industrial and agricultural soils, and long-term degradation of cementitious engineered barriers.

Todd A. McAlary Geosyntec Consultants, Inc., Guelph, Canada, tmcalary@geosyntec.com

Todd A. McAlary MSc, PEng, PG has a background in contaminant hydrogeology and chemistry, and has over 20 years of international experience in consulting and applied research. He is a Principal and the Vapor Intrusion Practice Leader at Geosyntec Consultants, Inc. and has been a member of the USEPA Expert Panel on Vapor Intrusion since 2000.

Mike J. McLaughlin CSIRO Land and Water/University of Adelaide, Glen Osmond, SA, Australia, Mike.McLaughlin@csiro.au

Prof. Dr. Mike J. McLaughlin is a Chief Research Scientist with CSIRO Land and Water and a Professor in the School of Earth and Environmental Sciences at The University of Adelaide, Adelaide, South Australia. His research interests are principally in the biogeochemistry of nutrients and pollutants in soils, food chain transfer of contaminants and the remediation of contaminated soils.

Helmut Meuser University of Applied Sciences, Osnabrück, Germany, h.meuser@fh-osnabrueck.de

Prof. Helmut Meuser PhD has a background in landscape planning and soil sciences and is working on soils of urban, industrial, traffic and mining soils. He has been with the University of Applied Sciences Osnabrück, Germany since 1997. His research projects are focused on assessment and clean-up of contaminated soils as well as functional soil estimation.

C. Paul Nathanail ARCADIS, Den Bosch, The Netherlands; University of Nottingham and Land Quality Management Ltd, Nottingham, UK, paul@lqm.co.uk

Prof. C. Paul Nathanail PhD, CGeol SiLC EuroGeol combines research and teaching at the University of Nottingham with consultancy practice through Land Quality Management Ltd. He directs a unique professional part time Masters programme in risk-based Contaminated Land Management and is a director of www.CABERNET.org.uk, Europe's sustainable Brownfield regeneration network.

Willie J.G.M. Peijnenburg National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands; University of Leiden, Leiden, The Netherlands, willie.peijnenburg@rivm.nl

Prof. Dr. Willie J.G.M. Peijnenburg is an organic chemist by training and his interests are in the area of fate and assessment of organic and inorganic pollutants. He is with the National Institute of Public Health and the Environment (RIVM), the Netherlands, since 1988 and is Professor in Environmental Toxicology and Biodiversity at Leiden University, the Netherlands, since 2009. His specific expertise is in the areas of bioavailability assessment, predictive ecotoxicology, degradation pathways, and fate and effect assessment of nanomaterials.

Robert Pentel GDF SUEZ, Paris, France, robert.pentel@gdfsuez.com

Robert Pentel MSc is involved in remediation and redevelopment of contaminated sites since 1998.

Anita Peter Institute for Geosciences, Christian-Albrechts-University, Ludewig-Meyn-Str. 10, 24118 Kiel, Germany, anita.peter@gpi.uni-kiel.de

Dr. Anita Peter has been teaching and carrying out research at the Institute of Geosciences at Kiel University as well as the Centre of Applied Geosciences at Tübingen University. For ten years she has been working on identifying and quantifying Natural Attenuation, focussing on processes simulation and the development of field concepts to prove Natural Attenuation.

Leo Posthuma National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands, leo.posthuma@rivm.nl

Leo Posthuma PhD has a background in ecology and ecotoxicology, especially on ecological effects of mixtures of contaminants in the field. He is main editor of a book on Species Sensitivity Distributions, senior project leader in ecology and ecotoxicology and Head of the Department of Soil at the National Institute of Public Health and the Environment (RIVM), the Netherlands.

Jeroen Provoost Flemish Institute for Technological Research (VITO), Mol, Belgium, Jeroen.Provoost@yahoo.co.uk

Jeroen Provoost MSc has a background in Environmental science and worked on contaminated land management since 1996. He worked three years in a consultancy firm, nine years at the Flemish Institute for Technological Research

(VITO) in Belgium and since 2008 for a European agency. His expertise is human health risk assessment, multi-media modelling and vapour intrusion.

Rene Rietra Wageningen University Research Centre, Wageningen, Alterra, The Netherlands, rene.rietra@wur.nl

Rene Rietra PhD has a background in soil chemistry and has been working on contaminated soils since 1999. He is with Wageningen University Research Centre (WUR, Alterra), the Netherlands, since 2002. His expertise is monitoring and risk assessment of diffuse soil pollution in agriculture.

Huub H.H.M. Rijnaarts Wageningen University, Wageningen, The Netherlands, Huub.Rijnaarts@wur.nl

Prof. Dr. Huub H.H.M. Rijnaarts has a background in environmental (bio)technology, has been working on contaminated soil, groundwater and sediments since 1988, and worked at TNO and Deltares (institute for delta technology) in the Netherlands. In 2009 he became professor at the department of Environmental Technology at Wageningen University, The Netherlands.

Massimo Rolle University of Tübingen, Tübingen, Germany, massimo.rolle@uni-tuebingen.de

Massimo Rolle PhD, PE has a background in environmental engineering and is working as a research fellow at the Center for Applied Geosciences, University of Tübingen (Germany). His expertise is modelling the fate and transport of contaminants in the subsurface.

Michiel Rutgers National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, michiel.rutgers@rivm.nl

Michiel Rutgers PhD has a background in microbiology and is working on contaminated sites at the National Institute of Public Health and the Environment (RIVM), the Netherlands, since 1990. His expertise is microbial physiology, soil ecology and environmental risk assessment.

Lida Schelwald-Van der Kley Envision-S b.v., Lelystad, The Netherlands, mail@schelwald.nl

Lida Schelwald-Van der Kley MSc has a background in biology and environmental studies. She is an independent consultant with ample experience in contaminated land management since 1987 (www.schelwald.nl/envision). She also acts as a Secretary to the NICOLE industry Subgroup. Her expertise is risk management and communication regarding the remediation of industrially contaminated sites.

Rosalind A. Schoof ENVIRON, Seattle, WA, USA, rschoof@environcorp.com

Rosalind A. Schoof PhD is a toxicologist certified by the American Board of Toxicology. For several decades she has been assessing human health risks from exposure to chemicals at contaminated sites. Dr. Schoof is recognized for developing methods for assessing the bioavailability of chemicals from soil. Currently she is working at ENVIRON.

Suresh Seetharam Belgian Nuclear Research Centre SCK•CEN, Mol, Belgium, sseethar@sckcen.be
Dr. Suresh Seetharam is a researcher at the Belgian Nuclear Research Centre SCK•CEN. He has nearly seven years experience in the development of numerical models for coupled thermal, hydraulic, chemical and mechanical behaviour in unsaturated soils. At SCK•CEN his work focuses on new developments in the field of durability assessment for cementitious engineered barriers.

Jiří Šimůnek University of California Riverside, Riverside, CA, USA, Jiri.Simunek@ucr.edu
Dr. Jiří Šimůnek is a Professor of Hydrology at the University of California Riverside. His expertise is in numerical modelling of subsurface water flow and solute transport processes. He has authored over 180 peer-reviewed publications. His HYDRUS models are used by virtually all scientists, students, and practitioners modelling water flow and chemical movement through soils.

Hans L.A. Slenders University of Nottingham and Land Quality Management Ltd, Nottingham, UK; Arcadis, Den Bosch, The Netherlands, h.slenders@arcadis.nl
Hans L.A. Slenders MSc civil engineering has been working in research and consultancy projects on contaminated soil and groundwater since 1989, and combines decision making with expertise of remedial techniques and soil processes. Within the consultancy and engineering firm ARCADIS he is responsible for innovation and development in the field of sustainable remediation and groundwater energy.

Erik Smolders Catholic University Leuven, Heverlee, Belgium, Erik.smolders@ees.kuleuven.be
Prof. Dr. Erik Smolders is professor of Soil Science at the Catholic University Leuven. His research deals with bioavailability of trace metals in soil, terrestrial ecotoxicology and risk assessment. He regularly serves as an adviser on soil contamination issues for government and industry.

Frank A. Swartjes National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands, frank.swartjes@rivm.nl
Frank Swartjes PhD with a Soil Science and Hydrology background, has been working in the field of risk-based contaminated site management since 1986. He is currently employed at the National Institute of Public Health and the Environment (RIVM) in The Netherlands.

Glenn W. Suter US Environmental Protection Agency, Cincinnati, OH, USA, suter.glenn@epa.gov

Dr. Glenn W. Suter II PhD is the Science Advisor in the U.S. Environmental Protection Agency's National Center for Environmental Assessment, Cincinnati, Ohio. He has experience in ecological risk assessment, ecological epidemiology and ecotoxicology.

Rudi Torfs Flemish Institute of Technological Research (VITO), Mol, Belgium, rudi.torfs@vito.be

Ir. Rudi Torfs MSc was trained as a physicist and an engineer, and has been working on environment and health related issues at the Flemish Institute for Technological Research (VITO) since 1997. He has collaborated in studies on health impact assessment and externalities of environmental pollution. Currently he is program manager of the research team on exposure and risk assessment modelling.

Stefan Trapp Department of Environmental Engineering, Technical University of Denmark, Lyngby, Copenhagen, Denmark, sttr@env.dtu.dk

Stefan Trapp studied geoecology, received a PhD in botany, made his doctorate in mathematics and is now associate professor for applied ecology at the Technical University of Denmark. Research areas are mathematical modelling and uptake into plants, applied in environmental engineering, risk assessment, phytoremediation and drug/pesticide design.

Robert H.M. Van de Graaff Van de Graaff & Associates, Melbourne, Australia, vdg.robert@vandegraaff-soilshorizons.com.au

Ir. Robert H.M. Van de Graaff PhD is Director of van de Graaff & Associates, a one-man consulting company (www.vandegraaff-soilshorizons.com.au), Melbourne, Australia. He was educated at Wageningen University, The Netherlands, and Cornell University, Ithaca. NY. He specialises in soil chemical aspects of treated wastewater irrigation, "detective" work on unexplained heavy metals in soils, as well as soil rehabilitation.

Martinus Th. Van Genuchten Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, rvangenuchten@yahoo.com

Dr. Martinus Th. Van Genuchten is a soil physicist, originally with the U.S. Salinity Laboratory, USDA-ARS, Riverside, CA and currently with the Federal University of Rio de Janeiro, Brazil. He has published widely on variably-saturated flow and contaminant transport processes and modelling in the subsurface. He is a recipient of the Soil Science Society of America's Don and Betty Kirkham Soil Physics Award, and fellow of the Soil Science Society of America, American Society of Agronomy, American Geophysical Union and American Association for the Advancement of Sciences.

Mirja Van Holderbeke Flemish Institute of Technological Research (VITO), Mol, Belgium, mirja.vanholderbeke@vito.be

Mirja Van Holderbeke PhD was trained as a chemist and has been a researcher at the Flemish Institute for Technological Research (VITO), since 1999. Her present expertise is situated in the area of human exposure assessment of environmental chemicals, human health risk assessment and policy studies with regard to environmental quality objectives.

Joop J. Vegter Vegter Advice, Amstelveen, The Netherlands, joopvegter@mac.com

Joop J. Vegter PhD has a background in biology. He has been a policy-maker and a scientific adviser for soil protection and contaminated sites at the Dutch environment ministry and is presently working as an independent consultant. He has been involved in international contaminated sites policy networks since 1994.

Martina G. Vijver University of Leiden, Leiden, The Netherlands, vijver@cml.leidenuniv.nl

Martina G. Vijver PhD has a background in environmental chemistry and ecotoxicology. She has been assistant professor at the Institute of Environmental Sciences (CML) with the University of Leiden since 2005. Her research focuses on the impact of chemical stressors on ecological systems, from a wide perspective.

Steve Wallace Secondsite Property Holdings Ltd., Altrincham, UK

Steve Wallace PhD key responsibilities include development of all aspects of sustainable environmental policies, strategies and operating models. He has been responsible for delivery of all aspects of National Grids contaminated land strategy in the UK, from developing 5 year business plans through to delivery of site specific risk management projects.

Joanna Wragg British Geological Survey, Keyworth, Nottingham, UK, jwrag@bgs.ac.uk

Joanna Wragg PhD, MRSC, CChem, CSci has a background in Analytical Chemistry and the bioaccessibility of soil contaminants with respect to human health, carrying out research on bioaccessibility and contaminant mobility within the Medical Geology and Abandoned Mines teams for the British Geological Survey. She is the Secretary of the Bioaccessibility Research Group of Europe (BARGE), www.bgs.ac.uk/barge.jwrag@bgs.ac.uk

Part I
Introduction

Chapter 1
Introduction to Contaminated Site Management

Frank A. Swartjes

Abstract Over thousands of years, contaminants have been added to the world's upper soil layers and have led to contamination of the soil and the groundwater. However, it was not until the late 1970s that several notorious cases of contaminated sites led to a sudden awareness to the general public. Today, in most developed countries, the number of potentially contaminated sites has grown to six or seven digits. This chapter describes the basic principles of contaminated site management. It focuses on risks and Risk Assessment, that is, quantifying the risks from contaminated sites on the basis of chance (exposure) and effects. This process is widely accepted today as offering the best balance between a sound scientific basis and practical implementation for appraisal of contaminated sites. Moreover, this chapter describes Risk Management, this is the process that brings contaminated sites back into beneficial use. The four major protection targets are human health, the soil ecosystem, the groundwater and Food Safety. Specific attention will be given in this chapter to a wide variety of topics including public and political awareness, soils, local and diffuse contaminated sites, contaminants, background concentrations, emissions to soil, site characterisation, land use, Soil Quality Standards, Brownfields, cost-benefit analyses, Risk Perception and Risk Communication, sustainability, and the actors involved in contaminated site management. Finally, several approaches to contaminated site assessment and management will be described, including the Fitness-for-Use approach, and Risk-based Land Management. In doing so, specific attention will be given to practical aspects such as effective use of financial resources and integration of contaminated site management (e.g., with regard to spatial planning, socio-cultural issues, economics and other factors).

F.A. Swartjes (✉)
National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: frank.swartjes@rivm.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_1,

Contents

1.1 Status of Contaminated Sites

1.1.1 History

1.1.1.1 Early Soil Contamination

Over thousands of years, since humans started mining for the iron-containing mineral hematite and later for malachite for copper production, potentially harmful chemical compounds (*contaminants*) have been added to the upper soil layers. And for centuries humans have dumped their waste materials into primitive waste dumps. However, large scale mining and, hence, large scale soil contamination, only came into existence in Europe, the USA and many other parts of the world in the 19th century. One phenomenon that sped up soil contamination was the Industrial Revolution which began in England and subsequently spread to several developed countries in Europe, the USA and Japan, from the turn of the 18th and 19th centuries. But the impact of the Industrial Revolution on contaminated sites was minor compared to the impact of the technological developments that took place mainly during the 20th century. These developments were characterized by a more than proportional increase in emissions of contaminants into the environment. Soil can often be considered as the ultimate sink for contaminants that enter the environment. As a consequence, emissions of contaminants to soil increased, for example, through the large-scale use of fertilizers, expansion of industrial production, the use of fossil fuels and, as an overall impact factor, a huge increase in population growth. It was not only the bulk rate of production of contaminants that significantly expanded. It was also the enormous increase in variety of types of chemical compounds that were produced for public or industrial use, or produced as a by-product, and eventually entered the environment and the soil.

The early examples of contaminated sites mainly resulted from delocalisation of contaminants, that is, metals in metal ores and crude oil from deeper soil layers to the upper soil layer. In the 20th century, however, an enormous variety of organic compounds, along with metal organic complexes, were produced out of existing less harmful compounds. Moreover, the soil was often intentionally used as a sink, for example, by primitive land filling or the release of contaminant holding fluids into the soil. Up to the 1970s, it was often believed that, very much like dumping waste water in a kitchen sink, superfluous materials simply disappeared towards some unknown destination where it was out of sight and, probably, would do no harm. A slightly more sophisticated approach towards the dumping of contaminant-holding materials was based on the belief that the soil-groundwater system was able 'to handle the burden', either by incorporating the contaminants in some kind of physical, chemical or biological cycling process, or simply by dilution. Although this latter approach included some arguments that we use in modern Risk Management procedures in regard to contaminated site management today, the power of the soil to 'clean' itself was far from being able to counterbalance the increasing contaminant load. Given the cost ratio between prevention measures and soil remediation, these approaches of dealing with contaminants must, in retrospect, be classified as immensely naive.

1.1.1.2 Public and Political Awareness

In the early 1970s, some soil protection-related policies came into existence in several countries. However, it was not until the late 1970s that several notorious cases of contaminated sites led to a sudden awareness among the general public and served as a loud alarm to decision makers. Those cases where contractors, generally without any bad intentions, had created situations in which humans came in close contact with notorious (carcinogenic!) contaminants in soil could especially count on intensive media attention.

In 1978 the *Love Canal disaster* became a national media event in the USA (Levine 1982). At the site of Love Canal, a neighbourhood near Niagara Falls in upper New York State, USA, a school and a number of residences had been built on a former landfill for chemical waste disposal, and thus sat directly on the dump site of thousands of tons of dangerous chemical wastes. The US Environmental Protection Agency (US EPA) discovered and reported on a disturbingly high rate of health afflictions for the residents, such as miscarriages and birth defects (Beck 1979). Although it was difficult to conclusively prove that the contaminants in the soil were the cause, recurring illnesses of the inhabitants and school employees were connected with the history of the site. In 1980, a state of emergency was declared and 700 families were evacuated.

In Europe in 1979, the site of *Lekkerkerk* in the Netherlands became an infamous national event. Again, a residential area had been built on a former waste dump which included chemical waste from the painting industry. Moreover, to prepare the site for the construction of a residential area, channels and ditches had been filled in with chemical waste-containing materials. The scandal started after a water pipe exploded because of the presence of aromatic contaminants in the soil. The specific known fact that benzene, a carcinogenic agent, was involved raised public concern. Nearly 300 families were evacuated, 1600 barrels of chemical waste were removed, and the soil under the residences was excavated.

Today, the Love Canal and Lekkerkerk cases still are often mentioned in introductions to reports on contaminated sites and in oral platform presentations at contaminated site management-related congresses.

1.1.2 The Present Situation

1.1.2.1 Extent of Soil Contamination

In the last two decades of the 20th century, the number of potentially contaminated sites grew in most developed countries to six or seven digits. During this period, most developed countries established monitoring systems for the purpose of assessing the extent of their contaminated sites. According to the European Environmental Agency (EEA) the number of contaminated sites requiring remediation in the EU member states was approximately 250,000 in 2007 (European Environmental Agency 2007). Today, it is expected that this number has grown significantly. According to the same source, potentially contaminating activities are estimated to have occurred at nearly 3 million sites (including the 250,000 sites already mentioned). In the European Union, 3.5 million sites are contaminated, affecting 231 million people and representing a market value of 57 billion Euros (Commission of the European Communities 2006). Soil contamination is one of the eight threats mentioned in the *EU Thematic Soil Strategy* (Commission of the European Communities 2006).

A contaminated soil map would roughly coincide with an anthropogenic map, since humans are generally recognized as the main polluters. Most of the contaminated sites are found in or close to cities.

In the present day, most countries have become aware of the huge practical, social and financial impact of contaminated sites.

1.1.2.2 Emissions to Soil

Emissions into the environment might occur through the air (atmospheric deposition) or directly (conscious or subconscious disposal). Examples of possible air emissions are:

- the deposition of Polycyclic Aromatic Hydrocarbons (PAHs) due to heating processes, often emitted tens or hundreds of kilometres away from the source (e.g., Ollivon et al. (2002), who measured substantial amounts of PAHs in atmospheric fallout (precipitation, gas phase and particulate matter), especially in winter, at an urban site in Paris, France);

- metal deposition from lead smelters (e.g., Salemaa et al. (2004), who measured elevated concentrations of metals in different plants, especially in bryophytes, near a copper-nickel smelter in Harjavalta, Finland);
- traffic (e.g., Hjortenkrans et al. (2006), who measured elevated concentrations of copper and antinomy due to decelerating activities, and lead and cadmium due to the combustion of petrol, in top soils in the south of Sweden);
- incineration activities (e.g., Schuhmacher et al. (2000), who measured elevated concentrations of Poly Chlorinated Dibenzo-p-Dioxins (PCDD) and Poly Chlorinated Dibenzo Furans (PCDF) in soil and vegetation in the vicinity of an old municipal solid-waste incinerator in Barcelona, Spain).

Metals and PAHs are known to be the most abundant and widespread contaminants worldwide.

Several other major contaminant sources are known to have contributed to large-scale soil contamination, such as coal combustion and mining activities. Another notorious source of contaminants, such as all kinds of petroleum hydrocarbons, PAHs, BETX (benzene, ethylbenzene, toluene and xylenes), methyl-*tert*-butylether (MTBE), and metals, is the oil industry, through oil exploration and production, refining and petro-chemical activities. In Mexico, for example, the number of reported hydrocarbon spills for the year 2000 exceeded 185 thousand, equivalent to 6252 tons (Iturbe et al. 2005).

In agricultural areas, contaminants have been introduced by using soils for waste-water filtering, applying sludge onto the soil, or applying ash from waste materials, used for liming (e.g., Pasquini and Alexander (2004), who demonstrated an increase of mainly lead through ash addition to soils on the Jos Plateau in Nigeria). Also the application of mine waste contributed to soil and groundwater contamination (e.g., Cobb et al. (2000), who demonstrated the presence of relatively high metal uptake of lettuce and radishes in soils mixed with mine waste in Bingham Creek, USA).

There also are many types of direct emissions. These direct emissions could follow from unintentional leakage (leaching) from artificial layers applied for soil surface-raising purposes, filling materials in construction works (dikes, roads) or from old waste disposal landfills, gas stations or storage tanks. Another example of an activity that generally caused contamination of soil and groundwater are gas works. See Fig. 1.1 for a picture of the former gas works ‘Delftse Wallen’ in Zoetermeer, the Netherlands, as an example, around 1908. At this site, heavy metals and PAHs were found in the upper soil layer, due to soil surface-raising activities, and aromatic contaminants, petroleum hydrocarbons and PAHs in the groundwater, in the late 20th century. Currently, a soil remediation is ongoing, the latest cost estimate is 2 million euro.

Another type of direct emissions to soil is spilling of contaminants during production processes, transport and storage. This often relates to (petro)chemical industrial sites (e.g., Nadal et al. (2004) who measured elevated concentrations

Fig. 1.1 The former gas works 'Delftse Wallen' in Zoetermeer, the Netherlands, around 1908, as an example of an activity that often caused contamination of soil and groundwater (photo: Historical Society 'Oud Soetermeer'; reproduced with permission)

of arsenic, chromium and vanadium around Tarragona, Spain, an area with an important number of petrochemical industries).

One specific kind of contaminant sources are activities at military training facilities. It generally includes a wide scale of polluting activities that might lead to human health and ecological risks (e.g., Teaf (1995), who dealt with human health and ecological risks at former military sites in the former Soviet Union). Another specific type of soil contamination arises from shooting ranges where lead bullets are deposited into the soils. In the state of Oregon, USA, for example, 211 active firing ranges exist (Darling and Thomas 2003). Soils in clay pigeon shooting ranges can also be seriously contaminated by heavy metals such as lead, antimony, nickel, zinc, manganese and cupper (Migliorini et al. 2004).

In some cases human or technical failure causes soil contamination. An example of this is a series of spilled mine tailing accidents. Since 1970, there have been 35 major mine-tailing dam failures reported (Macklin et al. 2003). One example of these is the collapse of a tailing dam in the Chenzhou lead-zinc mine in China in 1985, which led to the spread of huge amounts of mining waste spills onto farmland, followed by an emergency remediation procedure (Liu et al. 2005).

According to the European Environmental Agency, the most important sources for soil contamination in Europe, as an example, are industrial production and commercial service (41%), municipal waste treatment and disposal (15%), the oil industry (14%), industrial waste treatment and disposal (7.3%), storage (5.4%), power plants (3.9%), transport spills on land (2.1%), mining (1.4%), military (0.9%) and others (8.2%) (European Environmental Agency 2007).

A substantial part of the existing contaminated sites in developed countries are a legacy from the past. Today, however, it is widely recognized that the consequences of soil contamination were badly underestimated in the past. As a consequence, *prevention* of emissions into soil is a mandated practice in practically all developed countries today, for ethical, practical and financial reasons. Therefore, all kinds of prevention measures are being taken. Production processes, transport and storage follow strict regulations in order to avoid or minimize spilling. Gas stations have been equipped with liquid impermeable foundations. And waste gases have been filtered before leaving chimneys to avoid or minimize the exhaust of contaminants. And agricultural practices, including the processes that result in emissions into soil, are also strongly regulated today. An interesting tendency is the trend towards sustainable agriculture where the inputs into the soil are in equilibrium with the outputs.

1.1.3 Public Awareness

Out of the three major environmental compartments, that is, soil, water and air, *soil* probably is the least known and the least appreciated. Water, in particular surface water, is a highly visible and widely appreciated part of the landscape. The general public despises contaminated waters, for ethical and practical reasons. Since we are directly surrounded by it, the air compartment is as much appreciated as feared for. We need the air every minute of the day to survive and are alert for any disturbance in air quality. Moreover, we are very much aware of bad air quality, since we cannot help but see and smell it. Soil, on the contrary, is hardly visible. Except from an agricultural viewpoint, humans generally do not have a positive association with soil. To the general public, soil is often thought of as a dark place in which creepy organisms reside and in which we bury our dead.

Individuals who generally have a positive association with soil are people who grow crops, and realise the meaning of the soil in terms of habitat and nutrient source. These individuals can either be farmers who grow crops for commercial reasons, or individuals who grow crops for their own consumption (see Fig. 1.2, in which small allotments are shown in Jiangsu province, China, where the local population uses the spare bare surfaces for their food supply). An interesting initiative to make soils more palpable to the general public relates to the relationship between soils and art (Wessolek 2006). This movement reveals the beauty of soil profiles and of artworks that use soil materials or soil visions.

In the last few years the unfavourable view of soil has changed a little in Europe, the USA, Australia and Canada. Many people are now aware of the huge pressure that humans put on the environment, including the soil. Moreover, many environmentally conscious people read articles in the newspaper about the amazing performances of soils in, for example, organic biological agriculture practices. And since global warming is at the top of the political priority list, soils are recognized as a powerful CO_2 sink.

Not only scientists and regulators, but also the general public are all aware of the presence and consequences of soil contamination today. This awareness is still partly based on negative events in the 1970s and 1980s. Although the general public has good reason to be worried when their direct living conditions are impacted by

Fig. 1.2 Small allotments in Jiangsu province, China, used by the local population for their food supply (photo: F. Swartjes)

contaminants, this negative approach is sometimes exaggerated. There were several reasons for negative sentiments associated with those events of the 1970s and 1980s. First, humans were confronted with an unknown threat. Today we understand contaminated soils much better, and it is widely recognized that clear and objective information about a specific case, and the risks involved, is essential. Second, the citizens concerned felt betrayed by the government. When they bought real estate, no information was provided about the health risks they might be confronted with. In most countries today the transfer of property is accompanied by detailed information, often supported by computer systems that show the actual soil quality. In reality, in both the Love Canal and the Lekkerkerk cases, the government was as much overwhelmed by the phenomenon of soil contamination as were the citizens.

Following the Love Canal scandal, a clay lining was used to prevent further leaching of contaminants and a dirt cover was provided to prevent contact between the contaminants and humans. Unfortunately, these protective layers were damaged during construction work as a result of underestimating the threats involved. It is unlikely, however, that such technical mistakes, made in the early days of Risk Management, are still being made today.

1.1.4 The Contaminated Site Management Framework

1.1.4.1 Schematization

Several different frameworks for contaminated site management on the basis of risks are available. In Fig. 1.3 a schematization of the contaminated site management framework that is followed in this book, roughly in accordance with the highly influential report on Risk Assessment by the US National Research Council (US National Research Council 1983), has been illustrated in a simple graph.

The first step in this contaminated site management framework is *problem definition* (aka: *issue framing*). The second step is *Risk Assessment*, which

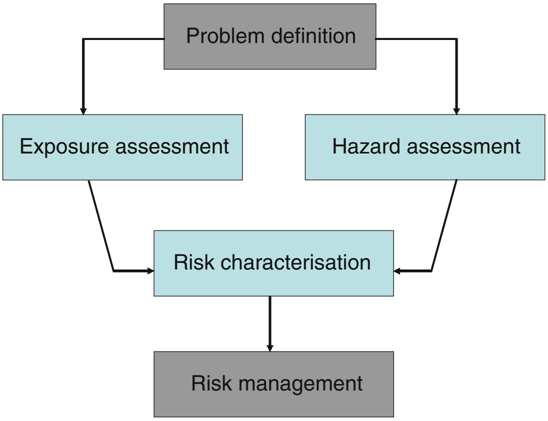


Fig. 1.3 The contaminated site management framework, as followed in this book (roughly in accordance with US National Research Council 1983). The *light-shaded boxes* relate to Risk Assessment

includes two different activities (see Fig. 1.3; the light-shaded boxes relate to Risk Assessment). These are the *Exposure Assessment* (aka: dose assessment) and the *Hazard Assessment* (aka: effect assessment), mainly used in Human Health Risk Assessment. The combination of the Exposure Assessment and the Hazard Assessment is called the *Risk Characterisation*, which results in an appraisal of the contaminated site. An alternative contaminated site management framework as used in the UK, including Risk Assessment, Risk Management but also economic and social issues, is given in Pollard et al. (2002a).

In several publications the phrase 'Risk Management' is used for the whole contaminated site management framework as shown in Fig. 1.3. In this book, however, *Risk Management* is the next logical step in the contaminated site management framework, following Risk Characterisation, in cases where this risk appraisal demonstrates the need for intervention, usually since the risk for a specific protection target is unacceptable. In this stage, solutions are sought for the purpose of bringing contaminated sites back into beneficial use, and are generally focused on risk reduction.

1.1.4.2 Problem Definition

The first step in a contaminated site management project is *Problem definition* or *issue framing*. In this step, the scope of the project needs to be clearly described. Moreover, the protection targets need to be defined. It is also very important to define the relevant time frame, since factors that impact risks will change over time.

Since regulators often have a profound impact on the initiation and performance of contaminated site management, it seems logical that they formulate the exact purpose of the contaminated site management project. Regulators also are responsible for defining the boundary conditions, for example, the required precaution/conservatism of the Risk Assessment for the given site. Therefore, intensive communication between regulators and scientists, and in fact between all stakeholders involved, must take place in the early stages of a contaminated site project.

1.1.4.3 Protection Targets

With regard to the definition of risk, which is a concept that denotes a potential negative impact on an asset, a relevant issue is determination of the nature and value of the asset. Since these assets are potentially negatively impacted, they need to be protected. Therefore, the term *protection targets* (aka: *receptors*) is often used.

With regard to contaminated sites, several protection targets have been recognised. The principle protection target with regard to contaminated sites, worldwide, is *human health*. More specifically, it is the physical health condition, not usually the mental health condition, of human beings that is considered. Alternatively, several risk-based quality assessment procedures use terms such as 'humans', 'human beings', or 'man' as protection targets, but they all refer to human health, that is, the state of physical health of human beings. There has, for decades, been an intensive debate on the extent of human health effects from contaminants in soil and groundwater. Although often overestimated, many studies have provided solid evidence that these effects are real. Beard and Australian Rural Health Research Collaboration (2005), for example, concluded that there is suggestive evidence for a role of exposure to DDT and DDE from soils with regard to pancreatic cancer, neuropsychological dysfunction, and reproductive outcomes. The relevant process with regard to the determination and often evaluation of risks is *Human Health Risk Assessment* (*HHRA*). Swartjes and Cornelis (Chapter 5 of this book) give a detailed overview of Human Health Risk Assessment. The subsequent chapters of this book (see Chapter 6 by Bierkens et al., Chapter 7 by Cave et al., Chapter 8 by Mclaughlin et al., Chapter 9 by Trapp and Legind, Chapter 10 by McAlary et al., Chapter 11 by Elert et al., this book) give details on the determination of Exposure Assessment, a crucial process in Human Health Risk Assessment. Langley (Chapter 12 of this book) describes the Hazard Assessment, another indispensable element in Human Health Risk Assessment.

A second protection target is the *soil ecosystem* (or Ecosystem health). Although not always appreciated to the extent it deserves, the soil ecosystem performs some immensely important tasks for humans (*Ecosystem Services*). Moreover, protecting the various species in soil contributes to the maintenance of Biodiversity. Only a few countries formally consider the soil ecosystem to be an protection target. Over the last few years, however, political and scientific interest in protection of the soil ecosystem has gained in importance, at least in Europe (Carlon and Swartjes 2007a). An enormous number of investigations have shown the adverse ecological effects of contaminants in soil. Nagy et al. (2004), for example, demonstrated the adverse effects of metals on nematodes in Hungarian soils. The relevant process with regard to the determination and often evaluation of risks is *Ecological Risk Assessment* (*ERA*). Swartjes et al. (Chapter 13 of this book) give an overview of Ecological Risk Assessment. The subsequent chapters of this book, these are Posthuma and Suter (Chapter 14 of this book) and Rutgers and Jensen (Chapter 15 of this book), describe Ecological Risk Assessment in more detail, from a generic and a site-specific perspective, respectively.

A third protection target is the *groundwater*. Juhler and Felding (2004), for example, demonstrated the presence of many organic contaminants in groundwater, including toluene, phenol, xylene, trichloromethane, benzene, dibuthylphthalate, 2,4-dichlorophenol, trichloromethane and pentachlorophenol, mainly originating from the upper soil layers, in 7671 groundwater samples collected from 1115 screens from the Danish National Groundwater Monitoring Program. Groundwater as a protection target has a special status, since groundwater is part of the soil as defined in the framework of this book. Moreover, the groundwater is both a protection target and a means of transport (*pathway*) for contaminant migration. The relevant process with regard to the determination and often evaluation of risks is called *Groundwater-related Risk Assessment* in this book. Swartjes and Grima (Chapter 17 of this book) give an overview of Groundwater-related Risk Assessment, considering groundwater both as a protection target and as a means of transport (pathway) of contaminants. The subsequent chapters (see Chapter 18 by Mallants et al., this book; and Chapter 19 by Rolle et al., this book) describe Groundwater-related Risk Assessment in more detail, that is, leaching of contaminants from soil into the groundwater and transport within the groundwater, respectively.

Finally, another important protection target is *Food Safety*. Generally speaking, this includes two different types of protection targets, namely, crops and animal products (meat, milk and eggs). An example is found in Yang et al. (2004), who evaluated the uptake of lead from soil into rice and meat, around a lead/zinc mine in Lechang, Guangdong Province, in China. The relevant process with regard to the determination and often evaluation of risks is called *Food Safety-related Risk Assessment* in this book. Specific elements with regard to Food Safety-related Risk Assessment, in particular to consumption of vegetables, are included in Mclaughlin et al. (Chapter 8 of this book), Trapp and Legind (Chapter 9 of this book) and Elert et al. (Chapter 11 of this book). The last chapter also includes a description of risk through consumption of animal products.

In addition to the protection targets, the level of protection also needs to be defined. The combination of protection target and protection level is often referred to as the ‘endpoint’.

The selection of appropriate protection targets and the level of protection in regulatory frameworks is primarily a policy decision. However, since the significance of protection targets and the levels of protection are often difficult to understand, policy decisions as to protection targets need to be supported by the scientific community.

1.1.4.4 Land Use

An important factor that affects both the risks and the degree to which those risks are evaluated, is the *land use* (often called: *function*) at a contaminated site. Generally speaking, the term land use applies to different categories that cover the main activity that is taking place on the site. Familiar land uses are *Residential land use*, *Industrial land use*, *Recreational land use, Children’s playgrounds, Infrastructural land use*, *Agricultural land use* and *Nature reserves*. Since sites with a similar land use can be used in quite different ways, the categorisation of land uses gives only a rough impression of the activities that are taking place at the site and the intensity of these activities. Therefore, land uses are sometimes subdivided, for example

Residential land use could be subdivided into ‘residential with garden’ (important with regard to exposure through vegetable consumption) and ‘residential without garden’ (no exposure through vegetable consumption). Moreover, several activities are covered by more than one land use. Housing, for example, is a prominent activity in Residential land use, but also occurs in Agricultural land use.

Risks for human health are strongly related to human behaviour. And human behaviour is highly impacted by the land use and activities taking place on the site. The degree to which risks are evaluated is mainly a policy decision. Generally speaking, the protection of human health risks warrants a greater weight in areas that are meant for human residence. It would be an option to give greater weight to human health protection in more densely populated areas, possibly with the weighting being proportional to the number of persons impacted, but this rarely occurs in existing contaminated site management frameworks.

In many countries with a high population density, land use changes in a relatively short term. The transformation from agricultural land to nature reserves and residential areas is especially common in many industrialised countries. In Europe, at least 2.8% of the land was subject to change in use between 1990 and 2000 (European Commission 2009) A change in land use has a profound effect on contact possibilities with the soil and on soil properties such as pH, organic matter dynamics (Römkens 1998) and, hence, mobility of contaminants and risks for humans, the ecosystem, the groundwater and Food Safety.

1.2 Soils and Sites

1.2.1 Soils

1.2.1.1 Definition

According to a broad definition, soil is the upper layer of the earth’s crust or, in geological terms, the exterior weathered part of the earth’s rocks. It has been formed out of rock material by physical, chemical and biological soil-forming processes over millions of years. Since climatical and geographical conditions varied over this long span of time, natural soils are typically characterised by a layered structure, that is, by the presence of *soil horizons*. According to a more popular definition, soil is the ‘skin of the earth, representing the inheritance of human history’. This metaphor reflects the vital nature of soil, while at the same time referring to the presence of historical soil contamination.

Most natural soils have a darker coloured upper layer, the A horizon, with higher organic matter levels. In many regions of the world, there is a loose organic matter layer of humus of a few centimetres on top of this A horizon, called an O horizon.

Soil, structured or non-structured, consists of three different phases, namely, a solid (mineral and organic materials), a liquid (*pore water*), and a gas phase (*soil gas*). Moreover, it contains plant roots and an enormous number of different organisms of a wide variety. Soil includes two different entities, namely, a water-unsaturated upper soil layer (*upper soil*) and a water-saturated groundwater zone. These two entities are separated by a groundwater table. Typically, the gas phase is absent in the water-saturated groundwater zone. From the perspective of groundwater subtraction for the drinking water supply, a water-saturated groundwater zone with high water volume and replenishing capacities is referred to as the *aquifer*.

Some definitions of ‘soil’ only refer to the water-unsaturated upper layer which, depending on the depth of the water table, implies a layer of several centimetres in swamp areas and up to hundreds of meters in arid regions of the world. Other definitions, for example, from an agricultural perspective, link the term ‘soil’ to that part of the earth’s crust that is actually used by humans. According to this definition soil usually includes the water-unsaturated upper soil layer and, often, the upper part of the water-saturated groundwater zone.

This book focuses on the impacts of contaminants. In this context, soil refers to that part of the earth’s crust in which contaminants reside that might impact one of the protection targets. Impacts from non-private water supply (from deeper groundwater or surface waters) are excluded from this book, since it is assumed that Waterworks sufficiently control the water quality. As a consequence, the scope of this book roughly coincides with the ‘agricultural definition’ of soil, that is, the water-unsaturated upper soil layer and the first tens of meters of the aquifer.

For practical reasons, the terminology followed in this book is linked with the most common terms used in soil policies and management of contaminated sites, see the schematisation of soil as defined from a wider perspective, in Fig. 1.4. In

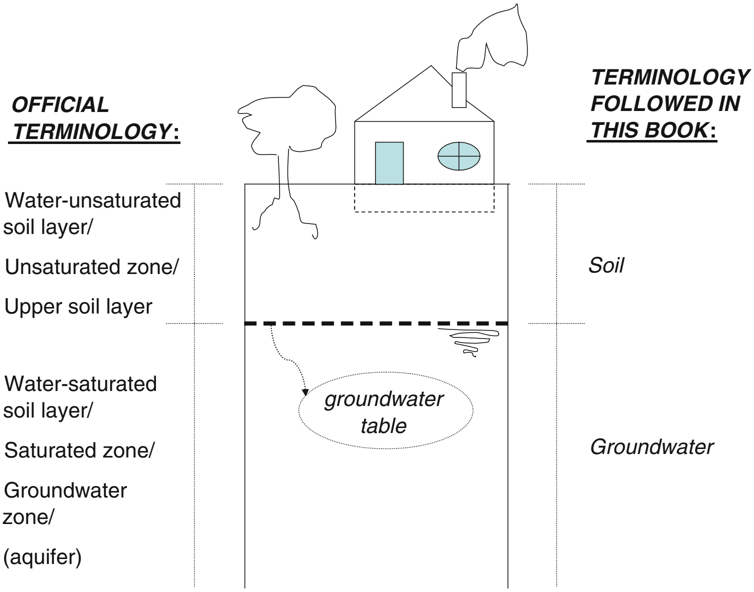


Fig. 1.4 Schematisation of soil

this much simpler terminology, the water-unsaturated layer above the groundwater table is often simply referred to as ‘soil’, while the water-saturated layer under the groundwater table is often called ‘groundwater’. Note that in this book ‘soil’ is sometimes also used in the wider definition, for example, in the context of ‘soil policy’, or ‘contaminated soils’. Clearly, both terms refer to both soil (in the more narrow definition of water-unsaturated upper layer) and groundwater.

An important difference between the upper soil and the aquifer is that groundwater is an important consumer product. This implies that clean groundwater has intrinsic value. From this perspective, the aquifer could be considered as a protection target. Soil, on the other hand, is not commonly used as a consumer product, but principally serves as an indispensable source for many useful products and activities.

It is important to realize that the separation between the upper soil layer and the aquifer, although from a scientific viewpoint convenient, is transient and partly artificial. With regard to the presence of contaminants, and the risks related to them, this separation is rather confusing. The reason for this is that contaminants migrate and do not necessarily belong to one of the two entities. Contaminants move throughout the soil-groundwater system, predominantly downwards, but sometimes also upwards and laterally, without acknowledging any borders between the soil and groundwater zones.

However, there are important differences between the water-unsaturated upper soil and the water-saturated aquifer. The upper soil, for example, enables the rapid transport of volatile contaminants via the gas phase. And the presence of a substantial amount of organic matter in the upper layer has a strong influence on the behaviour of contaminants and on the problems associated with these contaminants. Generally speaking, transport of water and contaminants is much faster in the groundwater than in the upper soil.

Another important difference between the upper soil and the deeper layers is the biological activity. Although organisms are found at every depth in the soil profile (see Chapter 13 by Swartjes et al., this book), the number of organisms is higher in the water-unsaturated upper soil, due to the presence of a gas phase. Within this upper soil layer, the number of organisms is even higher in the top of the soil, that is, the organic matter-rich layer that varies from a few centimetres up to several decimetres.

In many inhabited areas in the world the natural soil profiles are often disturbed. Many human activities, from the past and the present, are responsible for this feature, for example, (mechanical) digging activities in cities, tillage in agricultural areas or the addition to soils of foreign materials such as debris, stones, tar, and waste materials. Since most contaminated sites are within urban areas and disturbed soils are complex due to their heterogeneity, urban soil science is seen as a challenging, current frontier of soil science today (Norra 2006). In more extreme cases, whole new layers of soil material, mostly sand, sometimes clay, have been added onto the (natural) soil in many urban areas for infrastructural or filling purposes. Man-made soils, often with a high contribution of extraneous materials, are called *Technosols*. Generally speaking, soil structure is lacking in these artificial soil layers. In the lower parts of the Netherlands, for example, from the 16th century until

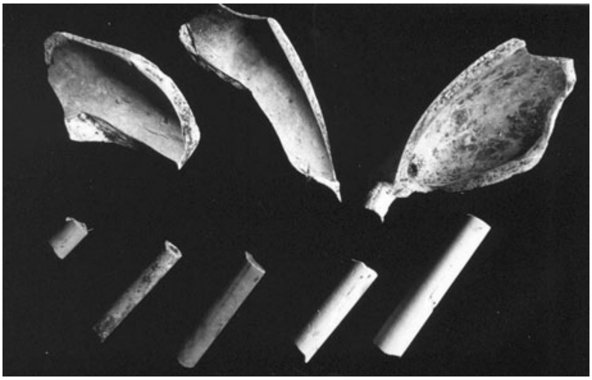


Fig. 1.5 Potsherds and pipe remains, found in the upper soil layers of the Central Western Peat area of the Netherlands, as evidence of layers of municipal waste of often several decimetres that have been brought onto the land from the 16th century until the 1940s (photo: M. Rutgers; reproduced with permission)

the 1940s, layers of municipal waste of often several decimetres have been brought onto the land in the Central Western Peat area, including potsherds and pipe remains (see Fig. 1.5), over areas as large as hundreds of square kilometres. Meuser and Blume (2001) describe the problematic classification of man-made soils, with regard to the anthropogenically influenced soils around the city of Osnabrück, Germany. Some of these soils contain coal and ore mining materials and waste materials. Meuser and Van de Graaff (Chapter 2 of this book) give a detailed description of the characteristics of and processes related to natural soils, urban soils and Technosols.

In fact, in all countries in the world all kind of waste materials have been brought onto the land over many centuries for the purpose of getting rid of these materials, whether or not in combination with soil improvement. After decades or centuries of evolution, these soils might have developed their own structure, for example, with an organic matter-rich upper layer. In any soil, whether disturbed or not, unconsolidated rock material on top of consolidated rock is found at greater depths. Typically, human activities that directly caused soil contamination also have resulted in disturbance of the original soil profile.

Unlike most bodies of water and air volumes, the soil is often privately owned. It is widely recognised that soil is a valuable and, at least on the time scale of decades, a non-renewable material. It serves several functions that are crucial for human survival such as crop production and as a supporter of buildings and humans themselves. Moreover, soil is the habitat and nutritional source for organisms.

1.2.2 Contaminated Sites

Contaminated sites are locations at which the soil and/or the groundwater are chemically polluted. In this context, a broad, three-dimensional definition is given to the concept 'contaminated site', including the soil (upper soil and aquifer) *underneath* the surface and the human occupation *on* the surface.

The focus of this book is on the threats for the four major protection targets; these are human health, the ecosystem, the groundwater and Food Safety, as related to contaminants present in the soil or the groundwater. From this perspective, the extent of the earth's crust that is relevant in the context of this book is limited to that part that impacts human health and the ecosystem, and the groundwater that is within human reach. More concretely, this primarily relates to the upper, unsaturated soil layer and the first tens meters of the groundwater layer. To a lesser extent, those groundwater layers will be considered from which groundwater is extracted (up to several hundreds of meters). Since this book relates to compounds that have adverse effects, these compounds are called *contaminants* throughout the book.

Two types of contaminated sites exist with regard to the extent and shape of the location that is contaminated, namely, *diffuse* and *local* contaminated sites. This extent and shape of a contaminated site is often dependent on the type of source that is responsible for the contamination. Generally speaking, atmospheric deposition and, to a lesser extent, large scale agricultural activities lead to diffuse contamination. Diffuse contaminated sites caused by atmospheric deposition are characterised by large contaminated areas. Often the contaminant concentration decreases along regular circles from the source, for example, in the case of lead smelters (e.g., Filzek et al. (2004), who measured the metal concentrations in soil along a transect from a smelter at Avonmouth, UK), where the concentration contours are possibly stretched according to the wind direction. Diffuse contaminated sites caused by agricultural practices generally are characterised by a relatively homogeneous contamination pattern. One specific version of diffuse contaminated sites is known as *ribbon contaminations*, for example, along roads or railroad tracks.

Locally contaminated sites vary in size from a small back yard of a few square metres to an industrial site of several tens of thousands of square metres. These locally contaminated sites generally are characterised by a heterogeneous contamination pattern, often with one or more cores (hotspots) of contamination, related to the source of the contamination. In many cases, the larger locally contaminated sites can be considered as a collection of smaller locally contaminated sites. There is no absolute definition of diffuse or local contaminated sites. Some sites have characteristics of diffuse and local contaminated sites combined, for example, in large diffusively contaminated sites with contaminant hotspots.

This book deals with contaminated sites, either diffusely contaminated or locally contaminated. However, since most chapters of this book deal with Risk Assessment tools that can be used for any type of contaminated site, this distinction is not always relevant.

In principle, the book does not implicitly deal with agriculturally managed sites. The reason for this is that agricultural activities often lead to a continuous supply of contaminants to the soil as part of agricultural business. This means that managing the contaminant status of agricultural sites, and the related risks, is a matter of balancing the inputs and outputs of contaminants. Pesticide application, for example, is focused on administering the applications needed for the goal to be reached (for example, prevention of crop diseases), while the load for soil and groundwater must be acceptable. The consequence is that the soil inputs from agriculture have been regulated in specific legislation in practically all countries in the world.

1.3 Contaminants

1.3.1 Terminology

No chemical substance leads to toxic effects by definition. Whether substances will cause toxic effects depends on the combination of exposure, the nature of the substance and the characteristics of the receptor (a human being or an organism). The overall exposure depends on the dose which humans or organisms are exposed to, the period over which this exposure takes place, the frequency of the exposure and the form (*speciation*) in which the chemical substance is available.

Several terms are used for the very generic term ‘chemical substance’ in contaminated site management. Often the term ‘compound’ is used. However, this term is considered too generic in the context of this book. Moreover, it literally does not cover all chemical substances in soil, since ‘compound’ refers to chemical substances that are composed of two or more elements, which means that pure metals are excluded. Alternatively, the term ‘compound of concern’ (or ‘chemical of concern’) (COC) is sometimes used. Of all the terms used, the word ‘pollutant’ evokes the most negative association, since definitions include adjectives such as ‘harmful’, ‘unsuitable’ or even ‘toxic’. The term ‘contaminant’ is used throughout this book, although it also has a negative ring, but this term does the best justice to the ‘potential’ aspect of causing adverse effects.

1.3.2 Daily Life

Potentially toxic chemical substances in the soil (*contaminants*) are part of our daily life. In modern times, humans and organisms are continuously exposed to a wide spectrum of contaminants. Humans are surrounded by all kinds of materials that contain a variety of potentially harmful chemicals, on a daily basis. Cloth, furniture, decorative objects or children’s toys, all contain chemical substances that are potentially toxic. Humans even eat and drink materials and inhale air that contains contaminants that are designated on several lists of Soil Quality Standards. Analogously, soil organisms are surrounded by all kinds of contaminants. They also feed on contaminant-holding materials.

Actually, humans have been in contact with contaminants since early human existence, due to the presence of metals in the soil, for example, or through PAHs from the burning of wood and roasting of meat. However, as long as humans lived in equilibrium with nature, exposure was limited and the threat to human health from contaminant exposure was generally negligible.

1.3.3 Categorisation

Since hundreds of thousands of contaminants are present in the environment, it is useful to categorise them. Several criteria can be used for this purpose, such as 'related production processes' (for example, heavy metals from zinc smelters, and cyanide from gas works), 'type of application' (for example, pesticides) or 'chemical or physical characterisation'. A systematic categorisation is given in Fig. 1.6.

A popular policy-related categorisation is given here that is partially based on chemical or physical characterisation and, hence, chemical properties. This results in the following six categories:

- metals and metalloids;
- Polycyclic Aromatic Hydrocarbons (PAHs);
- monocyclic aromatic contaminants;
- persistent organic pollutants (POPs);
- volatile organic contaminants (VOCs);
- other organochlorides.

Note that some of these categories overlap. In addition, three other useful categories can be added, based on 'frequency of occurrence of contaminants in soils'; these are:

- other inorganic contaminants (other than metals);
- petrolium hydrocarbons;
- asbestos.

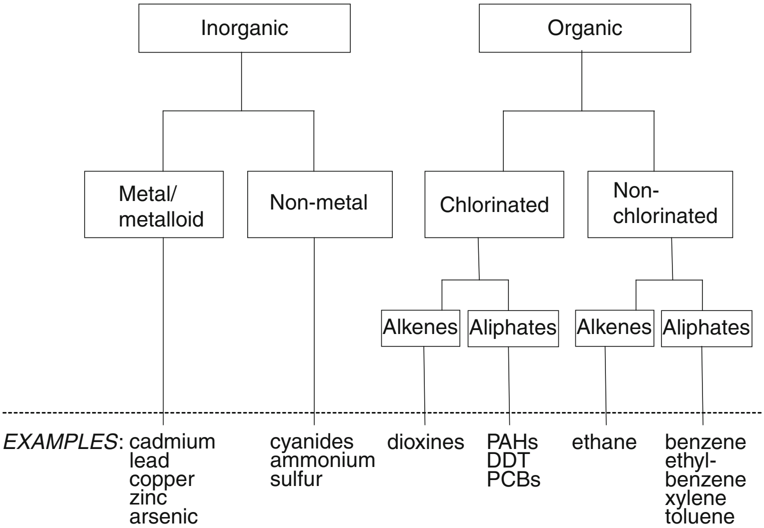


Fig. 1.6 A systematic categorisation of contaminants

1.3.3.1 Metals and Metalloids

‘Metals and metalloids’ is an important group of soil contaminants, since they are very often found in soils practically everywhere in the world. Metals (from the Greek word *metallon*) usually are characterised by a hard, malleable and shiny appearance, mostly solid at room temperature, with a high density and a high melting point and a good conductance of electricity and heat. Since metals readily lose electrons, they easily form positive ions (cations) in soils and, hence, have metallic bonds between metal atoms and ionic bonds with non-metals. Examples of the latter often found in soils are salts of metals and anions such as chloride (Cl^-) and carbonate ($CO_3{}^{2-}$). Different from most other contaminants, metals are elements included in the periodic table. Often the term ‘heavy metals’ is used for a sub-group of metals that are very often found at contaminated sites, although according to Duffus (2002) the term is controversial and archaic. Very important representatives with regard to soil contamination are cadmium (Cd), copper (Cu), lead (Pb), zinc (Zn), mercury (Hg), and selenium (Se).

Except for metals, this category of contaminants also includes the metalloids (or semi-metals), representatives of which, such as arsenic (As) and boron (B), are also often found in soils. These metalloids can be considered as transitional elements between metals and non-metals.

It is important to realise that metals and metalloids do not have a specific exposure affinity. The exposure depends on many factors, among them the form (speciation) in which the contaminant is available in the soil. This speciation follows from a combination of specific contaminant characteristics, soil properties and other chemical elements present. Many different speciations are found in soils for each metal. In this respect, a Soil Quality Standard for a specific metal could be considered as a lump standard for a whole group of chemicals which have the presence of the metal in common.

A general phenomenon in many residential areas is the presence of lead in soil, mainly due to the former use of tetraethyl lead in gasoline as an anti-knock agent (e.g., Wong and Xiang (2004), who measured elevated lead concentrations due to traffic activities in Hong Kong, China). This generally is a long-lasting major problem, since lead can result in retardation of the brain development of young children and is relatively immobile and will stay in the soil for decades or even centuries.

Another common problem is the presence of cadmium at agricultural sites or vegetable gardens (e.g., Wong et al. (2002) who measured enriched cadmium concentrations in crops, paddy and natural soils in the Pearl River Delta, one of the most developed regions in China; heavy metal enrichment was most significant in the crop soils, which might be attributed to the use of agrochemicals). Cadmium, which is often found in soils, due to atmospheric deposition from smelters or the application of fertilizers, is easily taken up by vegetables and can induce kidney dysfunction and several types of cancer at relatively low exposures.

1.3.3.2 Other Inorganic Contaminants (Other than Metals and Metalloids)

A specific case of other inorganic contaminants are the co-called *nutrients*. With regard to plant nutrition and, hence, soil contamination, the most relevant nutrients are the macro-nutrients nitrogen, phosphorus, potassium, and sulphur. These nutrients are needed in relatively large quantities in agricultural management, and are usually applied as nitrate, phosphate, potassium salts, and sulphate. Since this book does not primarily focus on agricultural practices, no further attention will be given to (the consequences of) nutrients.

A representative of the category of other inorganic contaminants often found in soils is cyanide, since cyanides are frequently found at former gas work sites, often in (inner) cities. Cyanide is a contaminant that contains a cyano group ($C\equiv N$) as a functional group, often found as the anion CN^-. Many organic contaminants feature cyanide as a functional group. Of the many kinds of cyanide contaminants, some are gases, while others are solids or liquids. Those that can release the cyanide ion CN^- are highly toxic.

Although it has been shown that rhizobacteria are cyanogenic (that is, able to synthesize cyanides), and hence negatively impact the seedling root growth of various plants (Kremer and Souissi 2001), most cyanide in soil has an anthropogenic origin. However, since cyanide is mainly present as iron cyanide complexes at gas work sites, the risk of effects on humans from exposure to cyanides often seems to be of minor relevance (Kjeldsen 1999).

1.3.3.3 Polycyclic Aromatic Hydrocarbons

Polycyclic Aromatic Hydrocarbons (PAHs) form also an extremely important group in regard to soil contamination, since they are among the most widespread contaminants found in soils, worldwide. They are characterised by a fusion of aromatic rings and do not contain many heteroatoms (atoms other than carbon or hydrogen). PAHs are primarily formed by incomplete combustion of carbon-containing fuels such as wood, coal, diesel, fat, tobacco, or incense (Fetzer 2000), and are concentrated in oil, tar and coal. Common PAHs in soil are naphthalene, phenanthrene, anthracene, fluoranthene, benzo[a]anthracene, benzo[k]fluoranthene indeno[1,2,3-cd]pyrene, benzo[g,h,i]perylene and benzo[a]pyrene. PAHs in soils might show local or diffuse (due to atmospheric deposition) contamination patterns. Different types of combustion yield different combinations of PAHs, both in terms of relative amounts of individual PAHs and with regard to the isomers that are produced.

Some PAH representatives are known or suspected to be carcinogenic, mutagenic, or teratogenic.

1.3.3.4 Monocyclic Aromatic Hydrocarbons

Some monocyclic aromatic hydrocarbons are frequently found in soil and groundwater. The representatives most often found are usually categorized as BTEX (benzene, toluene, ethylbenzene and xylenes). They were, or are, used on a large scale in cleaning applications such as degreasing.

Sort-term effects due to exposure of monocyclic aromatic hydrocarbons usually relate to skin and sensory irritation, as well as effects on the respiratory system and the central nervous system. Prolonged exposure to these contaminants also affects these organs as well as the kidney, liver and blood systems (Oregon Department of Human Services 1994). According to the US Environmental Protection Agency, there is sufficient evidence from both human epidemiological and animal studies to denote benzene as a human carcinogen. Workers exposed to high levels of benzene in occupational settings have been found to have an increased number of cases of leukaemia.

1.3.3.5 Persistent Organic Pollutants

Persistent Organic Pollutants (POPs) are organic contaminants that are resistant to chemical and biological degradation processes and to photolytic processes. For this reason they are capable of persisting in the environment and bioaccumulate in human and animal tissue. POPs are often halogenated, usually with chlorine. The United Nations Environment Programme Governing Council (GC) includes the following contaminants as POPs: aldrin, chlordane, DDT, dieldrin, endrin, heptachlor, hexachlorobenzene, mirex, polychlorinated biphenyls (PCBs), polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, toxaphene, certain brominated flame-retardants, some organ metallic contaminants such as tributyl tin (TBT), as well as some Polycyclic Aromatic Hydrocarbons (PAHs). Many POPs are pesticides which are banned in many countries, but still are found in soils and will reside in the soil for many more decades. In addition, POPs can originate from the production of solvents, polyvinyl chloride, and pharmaceuticals. Generally speaking, POPs have a high molecular mass and show a low water solubility, high lipid solubility and limited volatility.

According to the Stockholm Convention, POPs can lead to serious health effects, including certain cancers, birth defects, dysfunctional immune and reproductive systems, greater susceptibility to disease and diminished intelligence (Stockholm Convention 2009).

1.3.3.6 Volatile Organic Contaminants

Volatile organic contaminants (VOCs) are generally characterised by high enough vapour pressures under normal conditions to significantly vaporize. There is not one univocal exact definition for these contaminants. Under European law, the definition of VOCs is based on evaporation into the atmosphere, rather than reactivity. In the European Union Directive 2004/42/CE, for example, VOCs are defined as an ‘organic compound having an initial boiling point less than or equal to 250°C, measured at a standard atmospheric pressure of 101.3 kPa’ (European Union 2008). The US Environmental Protection Agency defines VOCs as ‘any compound of carbon, excluding carbon monoxide, carbon dioxide, carbonic acid, metallic carbides or carbonates, and ammonium carbonate, which participates in atmospheric photochemical reactions' (US EPA 2008). Infamous sources of VOCs are dry cleaning facilities. Other sources are paint, fabric softeners, petroleum fuels (e.g., gasoline), and crude oil. Moreover, several indoor sources are recognized, for example, photocopiers, carpet backings, and furniture. Widespread VOCs in soils include trichloroethylene, tetrachloroethylene, 1,1,1- trichloroethane, vinyl chloride, and, to a lesser extent, glycol ethers, hexane, formaldehyde, methyl bromide, methyl chloride and methyl ethyl ketone.

VOCs are readily soluble in fat. They may result in many different effects on human health, mainly after inhalation, ranging from dizziness, via narcotic effects, to neurotoxicological effects. Some agents of this group are carcinogenic, mutagenic or tetragenic.

1.3.3.7 Other Organochlorides

Organochlorides contain at least one chlorine atom. These chemicals are typically non-aqueous and are usually denser than water due to the presence of heavy chlorine atoms. The simplest forms of organochlorides are chlorinated hydrocarbons. These consist of simple hydrocarbons in which one or more hydrogen atoms have been replaced with chlorine. Many chlorinated hydrocarbons (e.g., dichloromethane, dichloroethene, trichloroethane, chloroform, and dioxins) are used as solvents. These solvents tend to be relatively non-polar and are therefore immiscible with water and effective in cleaning applications such as degreasing and dry cleaning. Other organochlorides are used as effective insecticides, such as DDT, heptachlor, endosulfan, chlordane, and pentachlorophenol. Polychlorinated biphenyls (PCBs) were once commonly used in electrical insulators and heat transfer agents. Their use has generally been phased out due to health concerns. Actually, BTEX (benzene, toluene, ethylbenzene and xylenes), here classified as monocyclic aromatic hydrocarbon, also could be included in this category.

Organochlorines generally affect the stomach, blood, liver, kidneys, and the nervous system.

1.3.3.8 Petroleum Hydrocarbons

Petroleum hydrocarbons (aka: petrol- or gasoline-related hydrocarbons; often called Total Petroleum Hydrocarbons or TPH) is a group of frequently found contaminants, which actually are complex mixtures of a whole spectrum of contaminants. These separate contaminants, which can add up to several hundred chemical compounds, mainly are hydrocarbons, both aliphatic and aromatic, and a whole spectrum of additives such as benzene, toluene, xylenes, naphthalene, and fluorene. No specific petroleum hydrocarbon mixture equals another existing petroleum hydrocarbon mixture. TPH compounds can affect the central nervous system, the blood, immune system, lungs, skin, and eyes or cause headaches, dizziness or a nerve disorder called 'peripheral neuropathy,' consisting of numbness in the feet and legs (ATSDR 2009). Several TPH compounds are (probably or possibly) carcinogenic.

For a long time, risk assessors seemed quite helpless with regard to the risk-based assessment of TPH. In some countries (e.g., the USA, the UK, Australia, and the Netherlands) expert judgement-based Soil Quality Standards have been implemented. An elegant approach for dealing with these complex mixtures was provided by Franken et al. (1999). They described a procedure for dealing with the human health risks of petroleum hydrocarbons, based on five groups of aliphatic hydrocarbons and five groups of aromatic hydrocarbons. Analogous to the US Total Petroleum Hydrocarbon Criteria Working Group, these hydrocarbon groups are characterized by a specific equivalent carbon number index range, representing equivalent boiling points.

An overview of the detection and remediation of soil and groundwater contaminated with petroleum products is given in Nadim et al. (2000).

1.3.3.9 Asbestos

The contaminant that provided the ultimate challenge for risk assessors, maybe even more than oil and petrol-like mixtures, is asbestos. Asbestos is also frequently found in soils. First of all, asbestos distinguishes itself from almost all other contaminants by the fact that asbestos is a mineral. It also is completely different from other contaminants in its behaviour: asbestos does not adsorb to soil particles and does not migrate through soils via the pore water, or soil gas. It is also not taken up by plants. The only pathway by which asbestos can give rise to adverse effects is by inhalation. These effects, although very serious (mesothelioma, that is, cancer of the pulmonary membrane and peritoneum, asbestosis, and increased risks for lung cancer), will reveal themselves over the longer term, that is, decades after exposure.

A concrete way of dealing with asbestos is described in Swartjes and Tromp (2008). They derived a Soil Quality Standard (Intervention Value) from measured data and described a tiered approach (as preferably used for other (composited) contaminants) to assess the site-specific risks of asbestos in soils. In the first step, measured asbestos concentrations in soil are compared with the Intervention Value of 100 mg/kg_{dw} asbestos equivalents (0.01% by weight). ‘Asbestos equivalents’ is the sum of the concentration of chrysotile asbestos (also serpentine asbestos or white asbestos) and 10 times the concentration of amphibole asbestos (other asbestos types), for both friable and bound asbestos. When this value is exceeded, a tiered approach is used for the determination of site-specific human health risks. A site-specific human risk is assumed, unless it can be proved otherwise (‘risk, unless. . ..’). The three tiers are as follows:

- Tier 1, Simple test: investigating the possibilities/likelihood of exposure;
- Tier 2, Determination of the *respirable fraction* in soil: investigating the possible site-specific exposure to humans, independent of the actual site use or site-specific elements, based on the determination of the respirable concentration of asbestos fibres in soil, in conformity with the Dutch standard NEN 5707.
- Tier 3, Measurement and testing of the concentration of asbestos fibres in outdoor and indoor air under standardised conditions.

1.3.4 Occurrence in Soils and Groundwater

Contaminants coming from various natural and anthropogenic sources may be found in soils. Contaminants enter the soil via emissions onto the soil surface, usually unintentionally (e.g., through atmospheric deposition, spills, etc.), sometimes intentionally (e.g., the use of metal-containing fertilizers, illegal dumping). In most cases, the contaminants migrate downwards. The velocity of migration varies greatly, depending on the type of contaminant, soil type, soil properties and climatic conditions. Ultimately, contaminants leach into the groundwater, where migration continues, both in vertical and in horizontal directions. As a consequence, contaminants are found in the entire depth range of a soil. However, since most soils have an upper layer with high organic matter content, which provides a high potential for adsorption of both inorganic and organic contaminants, generally the total soil concentration is higher 'in the first few decimetres' of the soil.

Since contaminant characteristics differ widely, the contamination profile vis-à-vis depth also differs. The shape of this profile is determined by the sorption, desorption and degradation potential of contaminants and of the physical characteristics (water flow transport) and physico-chemical characteristics (sorption and desorption) of the soil. Generally speaking, immobile contaminants have a higher ratio between the concentration in the solid phase of the soil and the concentration in the pore water or the groundwater than do mobile contaminants. However, all contaminants have higher concentrations in the solid phase of the soil and, hence, higher total soil concentrations in the upper soil, rich in organic matter and in clay horizons (mainly metals).

For practical reasons the concentration of contaminants in soil is usually expressed by weight of contaminants per unit weight of dry soil (kilogram), while the concentration of contaminants in groundwater commonly is expressed by weight of contaminants per unit of volume (l). Since this proved to result in the most convenient figures, the contaminant weight is usually expressed in milligram (mg) for soil and in microgram (μg) for groundwater. In summary, the concentrations are expressed as $\mathrm{mg_{contaminant}/kg_{soil,\ dry\ weight}}$ for soil and in $\mathrm{\mu g_{contaminant}/l_{groundwater}}$ for groundwater, most often shortened to $\mathrm{mg/kg_{dry\ weight}}$ (or mg/kg), and μg/l, respectively.

1.3.5 Mixtures of Contaminants

In the great majority of contaminated sites, more than one contaminant is found in soil or groundwater. There are two reasons for this. First, most materials from which contaminants originate contain more than one contaminant. Metal ores, for example, often contain several metals which may be simultaneously released from metallurgic industrial processes. Similarly, in most activities or processes, where contaminants are released into the environment, several contaminants are involved. One example of this is a dry cleaning facility, where several chlorinated hydrocarbons are simultaneously used. From both these examples it can be concluded that the same combinations of contaminants are often found in soils and groundwater. The risk assessor needs to use this information in order to investigate the site for the whole contaminant mixture. The exact composition of contaminants in soils and groundwater, however, may differ. By incomplete combustion of different organic materials, for example, different mixtures of PAHs are produced, depending on the type of organic material and combustion characteristics such as temperature.

Second, specific sites, for example sites just outside the city limits of several of the larger cities around the turn of the nineteenth and twentieth centuries, lent themselves to several soil contaminating activities. At these sites, an often incoherent cocktail of contaminants is present. For these sites, it is more difficult to determine which contaminants to search for.

1.3.6 Scope of This Book

A specific class of potentially harmful contaminants found in soil consists of *radioactive contaminants* (e.g., Callahan et al. (2004), who evaluated the human health risks due to the presence of depleted uranium at a military training site in the USA). Since radioactive substances are of a different nature and require a different kind of Risk Assessment, these contaminants do not fall within the scope of this book. For the same reason *endocrine disruptors* (aka: 'hormonally active agents'; see Lintelmann et al. (2003), who provided an overview of the biochemical and biological background of endocrine disrupters in the environment) are not considered in the scope of this book. Furthermore, no attention will be paid in this book to the *microbial contaminants*, mainly relevant in groundwater, that originate from both human and animal faeces via sewer leaks, septic tanks and manure disposal, although these are of great concern for human health (e.g., Celico et al. (2004), who found several microbial contaminants, related to pasture and/or manure spreading, in different carbonate aquifers of southern Italy).

Recently, there has been much attention paid to the impact of *nanoparticles* in the environment. Since these nanoparticles are central to many natural processes in soil and groundwater and in human physiology, they are a potential threat to the soil ecosystem and human health. Given the limited scope of their use, it is currently unlikely that they pose a substantial risk to the soil ecosystem and human health (Colvin 2003). However, since the widespread use of nanomaterials will result in higher concentrations in soils, the future impact is unknown. For the same reasons that radioactive contaminants and endocrine disruptors require a different kind of Risk Assessment, nanoparticles are not considered in the scope of this book.

1.4 Site Characterisation

Site characterisation is an essential step in identifying contaminated sites and, in the steps that follow, contaminated site management. Therefore, each project involving contaminated sites needs to begin with a preliminary study of the site under investigation. This study includes information on the layout of the site (the presence of buildings, sealed surfaces, bare surfaces, vegetation, (micro)relief, the presence of soil-foreign materials, elements that relate to former activities) and a detailed evaluation of the history of the site (which activities might have been responsible for which contaminants, at which spots on the site). Some sites evoke a clear suspicion of being contaminated, while other sites have a rather innocent appearance with regard to soil contamination. See Fig. 1.7, as an example, in which a suspicious site, based on the presence of drums (upper photo), and a site that seems above suspicion (lower photo), both in the Silvermines area in Tipperary county, Ireland, contaminated with several metals due to former mining activities.

A *site visit*, including a so-called *organoleptic investigation* ('looking and smelling'), is an essential activity at this stage of the project. The evaluation of the history of the site might include a visit to the municipal archives and the historical records found in the library. Moreover, interviews with former workers or inhabitants might be helpful. A map of the site might also help in the interpretation, and digital photos will support the memory of the risk assessor. This preliminary study should result in a hypothesis about the type of contaminants and the spots where these contaminants can be present.

Obviously, samples need to be taken and analysed in order to determine the concentrations in soil and groundwater. If no information about the possible contaminants is known, the samples can be analysed for a group of 'frequently found contaminants'. This group differs for soil and groundwater, since the more immobile contaminants are often found in soils, while the more mobile contaminants usually reside in the groundwater. Several countries have defined standard groups of frequently found contaminants, often formalised in protocols. In the Dutch NEN 5740 protocol, for example, a standard series of contaminants has been defined that must be determined when there is no information about the possible contaminants present (NEN 2009). The selected contaminants differ for:

Fig. 1.7 A suspicious site (**a**) and a site that seems above suspicion (**b**), in the Silvermines area in Tipperary county, Ireland, contaminated with several metals due to former mining activities (photo: F. Swartjes)

- soil: barium (Ba), cadmium (Cd), cobalt (Co), cupper (Cu), mercury (Hg), lead (Pb), nickel (Ni), selenium (Se), vanadium (V), zinc (Zn), chloride (Cl), mineral oil, sum of EOX (Extractable Organic Halogens) and the sum of PAHs (10 specified representatives);
- and groundwater: cadmium (Cd), cobalt (Co), cupper (Cu), mercury (Hg), lead (Pb), nickel (Ni), selenium (Se), vanadium (V), zinc (Zn), chloride (Cl) , mineral oil, naphthalene, some specified volatile aromatic hydrocarbons (including BTEX), and some specified volatile halogenated hydrocarbons.

Since sampling and laboratory analysis are relatively expensive, generally there is lack of data. Multivariate and geostatistical tools can support the characterisation of a site, e.g., Carlon et al. (2000) who extracted additional PAH concentrations by Kriging interpolation of spatial data and Principle Component Analyses (PCA), at an industrial site close to Parma, Italy.

Generally, two important decisions need to be taken, which require a combination of science and pragmatism. First, the number of samples needs to be determined. Often, the number of samples that results from a pure statistical analysis is too costly. Therefore, statistics need to be combined with pragmatism. In Lamé (Chapter 3 of this book) the procedure for sampling has been described, primarily from a practical perspective. In Brus (Chapter 4 of this book), this procedure is approached from a statistical perspective.

Second, a decision needs to be taken about the construction of *composite soil samples*; these are lumped samples through mixing of separate samples. Obviously, a chemical analysis of composite samples is factors cheaper than a sampling of the separate samples. In case the composite samples do not provide enough information for a well-founded risk appraisal, appropriate separate samples could be analysed at a later stage. The decisions pertaining to the number of individual samples and composite samples depend on the degree of heterogeneity of the contaminant (and of some important soil characteristics such as pH, organic matter content) in the soil and groundwater.

Also the (statistical) interpretation of the measured concentrations is important. Altfelder et al. (2002), for example, showed that part of the area that may be declared safe based on merely kriged estimates can actually exceed the German limit values by a probability of up to 50%. Millis et al. (2004) showed for lettuce (variety Crispino) that variation in plant-scale heterogeneity of cadmium in soil affects bioavailability and hence the concentration factors plant-soil by a factor of two.

1.5 Risk Assessment

1.5.1 Principles

A measured concentration in soil or groundwater is a rather vague criterion with regard to determine possible associated problems. The simple purpose of Risk Assessment is to transfer this measured concentration into a more manageable appraisal of the status of the contaminated site in terms of risks for one of the protection targets (human health, the ecosystem, groundwater or Food Safety). Let's consider, as an example, the appraisal of a PCB soil concentration of 1 mg/kg_{dw} soil. Without Risk Assessment, it is extremely difficult to give an objective and useful opinion about this measured concentration. To some, it might be a non-problem, since 1 mg/kg_{dw} soil implies one in a million, and that seems very low when compared with some (undefined) standard of high and low. But another person might approach the case differently, that is, by noting that 1 mg PCBs equals about 1.8. 10^{18} molecules.[1] Such a high amount of a contaminant that is able to impact the immune, hormone, nervous, and enzyme systems is associated with serious health problems, again by comparing it with some (undefined) standard of high and low. But obviously both positions are not very useful, since the numbers do not tell anything about the magnitude of the problem. And it is exactly that, an estimate of the magnitude of the problem, which is the purpose of Risk Assessment.

1.5.2 The Concept of Risk

Risk is a concept that denotes a *potential negative impact to an asset*. There must be a source for this potential negative impact, and this is generally called a hazard. With regard to contaminated sites, the hazards are the adverse effects on human health from contaminants in the soil or groundwater.

Many authors describe the magnitude of a risk in terms of probability (or change, or frequency) and effect (harm). Since a doubling of the probability of a negative impact on an asset often is judged similar to the doubling of the effect, risk is often described as the multiplication of probability and effect. The determination and often the evaluation of risks are called *Risk Assessment* and helps in making transparent, rational, and defendable decisions.

With regard to the seriousness of an effect, it is very important if, and if so, to what extent, one can influence the probability of a negative impact on an asset. In this respect, it is useful to distinguish between a risk that humans deliberately take, for example, the risk of getting lung cancer from smoking (a *voluntary risk*), and a risk that is beyond human control, for example, the risk of a natural catastrophe (an *imposed risk*). Humans can control voluntary risks, for example, by reducing the number of cigarettes they smoke. Imposed risks, on the contrary, are not or are difficult to manage. At best, if one is prepared to take extreme measures that often impact one's personal circumstances, some risks can be reduced, for example, by moving to a place on the globe where the chance of natural catastrophes is relatively low.

Risk, both voluntary and imposed risks, relates to a concept we deal with on a daily basis. Some examples of familiar voluntary risks, with human health as the asset that can be negatively impacted, relate to the consumption of alcohol-containing drinks, going out in traffic, and engaging in sport activities where injuries

---

[1] Assuming an average average molecular weight for PCBs of 327 g/mol.

to joints are possible. Clearly, we are prepared to take some of these risks, since these activities provide us with evident advantages. Some examples of imposed risks are living in polder regions beneath sea level, protected by dunes and dikes, the threat of natural catastrophes in hurricane-prone areas of the world and of terrorist attacks.

Risks from contaminated sites typically are in the category of imposed risks: humans can only avoid, or at best reduce the risks, by adapting their lifestyle or, in the most extreme case, by moving to a place where the soil has not been impacted by soil contamination. From the perspective of Risk Assessment, it is very important to realise that risk is not necessarily a bad thing. Being at risk is part of our life; many risks will never impact human health, and there are many risks we are not even aware of.

When contaminants are present in soil there is a risk by definition, since there is ‘a probability’ (the chance, although miniscule at very low concentrations, that human beings or soil organisms are exposed to these contaminants) and ‘an effect’ (impact on human health or ecosystem health, when contaminants indeed intrude on humans or soil organisms). The whole idea behind Risk Assessment is not to find out whether there are risks, but to investigate whether or not the risks are *acceptable*. Nevertheless, risk assessors often use the phrase ‘no risks’ when in fact ‘no unacceptable risks’, or more concretely ‘acceptable risks’ is meant. Often, regulators or stakeholders seduce the risk assessor into using the ‘no risks’ qualification, since the ‘no unacceptable risk’ qualification (double denial), or ‘acceptable risks’, is more difficult to interpret and often provokes further discussion.

Risk Assessment can also be used to support the optimal allocation of financial resources in contaminated sites projects.

1.5.3 Procedure

Risk Assessment (aka: Risk Analyses) is a process which serves the purpose of examining risks and, when possible, quantifying risks. It is an old concept. Risk Assessment can almost be considered as a science unto itself. It is used in widely differing disciplines such as environmental engineering, the design of building constructions, financial impact assessments or in the military. Analogous to risks, we are dealing with Risk Assessment daily, mostly without realising it. When a person crosses the road, for example, that person makes a judgement on the chance of being hit by a passing vehicle and the following consequences.

The Risk Assessment framework is illustrated by the light-shaded boxes in the contaminated site management framework, in Fig. 1.3. The first step in the Risk Assessment framework includes two different activities, namely, the Exposure Assessment (aka: dose assessment) and the Hazard Assessment (aka: effect assessment), mainly used in Human Health Risk Assessment. Conventionally, the determination of exposure is performed for human beings or larger animals, and not so much for smaller soil-dwelling organisms. Ideally, the amount of a contaminant that reaches the blood stream (in case of systemic effects, i.e., related to effects in the whole body after systemic circulation and, hence, absorption and distribution in the body) or target organs (in case of local effects, i.e., related to effects to specific organs at the place of contact or intake) is determined. This amount is expressed as mass contaminant, per body weight mass, per unit of time (mg/kg$_{\text{body weight}}$/day), that is, as the *internal* exposure. Target organs are the organs that could be adversely affected by specific contaminants. However, in most cases the *external* exposure is calculated, that is, the amount of a contaminant that reaches the human body or the organism.

With regard to the determination of human exposure, multimedia calculations are usually combined with the calculation of human exposure in so-called exposure models. A more detailed introduction to Human Health Risk Assessment is given in Swartjes and Cornelis (Chapter 5 of this book). A quantitative determination of human exposure is described in detail in Chapters 7, 8, 9, 10, and 11 of this book.

An important first step in Exposure Assessment is to determine the *representative soil concentration*. The representative soil concentration, and hence, the procedure for soil and groundwater sampling, is dependent on the purpose of the Risk Assessment, the site characteristics and the exposure pathways that are the most relevant. Smart choices need to be made for location and depth of the samples, construction of the composite samples, and contaminants that must be analysed (see Section 1.4). A site visit and historical survey can be important activities in support of the chosen sampling strategy.

Hazard Assessment includes two steps, namely, *Hazard Identification* and *Hazard Characterisation* (IPCS 2004). Hazard Identification focuses on the possible effects of specific contaminants and the time frame for which these effects occur. Subsequently, the Hazard Characterisation results in a dose-response assessment, relating exposure to effects, and is the basis for the determination of *Critical Exposure* (aka: Reference Dose).

The combination of Exposure Assessment and Hazard Assessment, the second step of the Risk Assessment framework, is called *Risk Characterisation*. When translated in objective terms, the Risk Characterisation results in a *Risk Index*. This is the ratio between *actual exposure* and Critical Exposure with regard to Human Health Risk Assessment, or the ratio between the actual concentration and acceptable concentration in the soil with regard to Ecological Risk Assessment, respectively. In Ecological Risk Assessments often the PEC/NEC (Predicted Effect Concentration/No Effect Concentration) ratio is used for one organism, several organisms or the whole ecosystem, with the same goal in mind. Dawson et al. (2007), as another example, established a Biological Soil Quality Index to help visualize significant differences in hydrocarbon-polluted soils.

The actual performance of Risk Assessment is generally supported by appropriate *Risk Assessment tools*. In Swartjes et al. (2009) a Risk Assessment tool is defined as any instrument that can contribute to the determination of risks at a contaminated site. A Risk Assessment tool can be an equation, a description, a database, a model, an instrument, a protocol, or a table. A combination of selected Risk Assessment tools is called a *Toolbox*. Such a Toolbox does not include policy points of view. The Toolbox for the determination of human health risks, for example, may include algorithms for the calculation of exposure through different pathways, a measurement protocol for the determination of the indoor air concentration, a table with Critical Exposure values, and many more Risk Assessment tools. A Decision Support System (DSS) does include selected Risk Assessment tools, but also policy points of view. A DDS could for example be used to calculate national Soil Quality Standards on the basis of several selected Risk Assessment tools and national policy points of view.

It is often said that Risk Assessment is an objective process and that scientists need to operate independently of the interests of any stakeholder. To a certain extent this is truth, since scientific independence is the key to an objective risk qualification. This independent position, however, certainly does not justify a strict 'no communication policy'. The reason for this is that Risk Assessment includes several policy decisions, for example, as to the degree of conservatism and the required level of protection of human health, the soil ecosystem, the groundwater or agricultural products.

The independent status of scientist will not be affected by the adaptation of specific political boundary conditions, as long as it is made transparent what these boundary conditions are. Risk assessors can do an excellent and objective job when they, for example, commit themselves to the political boundary condition that a Risk Assessment for an industrial site should focus on 'average adult workers' and not relate to children or other sensitive groups. Again, it is important to make these boundary conditions and, hence, the validity range of the conclusions from the Risk Assessment, transparent. Therefore, this political boundary condition needs to be clearly described in the Risk Assessment report. This enables regulators to guarantee the safety of these sensitive groups, for example, by fencing of the site with anti-trespassing controls in order to protect children as in the case just mentioned.

In the USA, there is a standardised procedure for performing Human Health and Ecological Risk Assessment called *Risk-Based Corrective Action (RBCA)*. A summary of this three-tiered Risk Assessment procedure is given in ASTM (2009).

It is generally acknowledged that the total concentration is not the optimal measurement with regard to risk to the soil ecosystem, the groundwater and, to a lesser extent, for human health. Especially exposure and leaching are strongly related to an 'effective' fraction of the contaminant in soil, this is, the *bioavailable fraction* with regard to ecological risks and a specific part of human health risks (like the risk through vegetable consumption) and the *available fraction* with regard to leaching from the upper soil into the groundwater. In Mallants et al. (Chapter 18 of this book) a detailed description of the leaching process is given. The (bio)available relevant fraction depends on the type of organism and, last but not least, the relevant timeframe.

An enormous number of papers have been written on calculating and measuring bioavailability, in particular with regard to metals. An example is given in Alvarenga et al. (2008), who determined two bioavailable metal fractions, that is, a 'mobile fraction' and a 'mobilisable fraction' using a sequential extraction, with the purpose to assess the risks in an acid metal-contaminated soil from the Aljustrel mining area in Southwest Portugal, in the Iberian Pyrite Belt. An example with regard to plant uptake is given in Kalis et al. (2007), who described a procedure for assessing metal uptake by *Lolium perenne.* To this purpose they used a four-step approach, starting with the total metal content in soil, including the calculation of the concentration in the pore water, the metal concentration adsorbed to the root surface, the metal contents in the roots and the metal contents in the shoots.

In Hodson et al. (Chapter 16 of this book) a detailed description of bioavailability is given.

1.5.4 Reliability

1.5.4.1 Uncertainties and Variability

It is generally acknowledged that Risk Assessment, although it is said to be an objective process, is also an unreliable process (e.g., Ferguson et al. 1998). There are several reasons for this. First, Risk Assessment includes many parameters and equations that have large uncertainties and variability. Uncertainty is the variation in these Risk Assessment tools due to lack of knowledge or to lack of scientific consensus. Variability is the variation due to spatial and temporal variations. The large variability is explained by the heterogeneous nature of soil and the large differences in human characteristics and behaviour among individuals. The large uncertainties, often found in exact sciences, are mostly related to the transfer of contaminants from soil into contact media, and biokinetic fate and transport processes in the human body (Human Health Risk Assessment and Food Safety), the immensely complex functioning of the soil ecosystem with mutual interactions between many different organisms and the soil properties (Ecological Risk Assessment), and transport processes in soil and aquifers (Groundwater-related Risk Assessment). Every variability found in the Risk Assessment tools is interwoven with uncertainty, while uncertainty does not necessarily go together with variability. Examples of equations that are characterised by both large uncertainties and variability are transport processes of contaminants through the aquifer, pore water and soil gas, and the equations that describe bioavailability in soil. A parameter that mainly has large variability is, for example, the fraction of total vegetable consumption that humans grow in their own garden. For a specific case this fraction can be accurately estimated, but for the use of a generic value for the derivation of Soil Quality Standards there is a huge variation between sites, and most definitely for bigger geographical entities. Generally speaking, patterns that describe human behaviour, and the behaviour and composition of soil ecosystems show a wide variation in time and space.

Second, Risk Assessment includes a whole chain of calculations and measurements, which means that small uncertainties in an earlier step (e.g., in the sampling strategy) might add up to large uncertainties in the final step (e.g., the risk characterisation). Third, several elements in Risk Assessment require a subjective judgment, which means that quite a number of uncertainties are involved with the sometimes arbitrary choices of the risk assessor.

It must be realised that measurements, although the general belief is that these are much more accurate than calculations, are also often characterised by limited reliability. Nevertheless, in some specific cases, the reliability could indeed be improved by including measurements in the calculations, namely, measurements of ‘supportive input parameters’ or the concentrations in contact media. By measuring the organic matter content of soil (a ‘supportive input parameter’), for example, clearly the reliability of a site-specific Risk Assessment can be improved, as compared to an assessment based on an average organic matter content of a specific region (elimination of variability).

A specific type of uncertainty relates to the lack of clear definitions of political boundary conditions or the wrong interpretation of these political boundary conditions by scientists. When, for example, the degree of precaution (e.g., whether the average human being or the great majority of human beings must be protected) has not been clearly defined, or is incorrectly interpreted, the input parameter identification of exposure parameters could take on an arbitrary character.

1.5.4.2 Dealing with Uncertainties and Variability

As was mentioned in Section 1.5.4.1, each input parameter for Human health, Ecological, and Groundwater-related Risk Assessment is characterised by uncertainty and variability. Nevertheless for many Risk Assessment applications it is useful to represent the input parameter by one single value. Options that are mostly used, depending on the type of Risk Assessment, the purpose of the Risk Assessment, and the possible political boundary conditions, are based on a single value of the *central tendency* or some kind of worst-case estimate. Most often, a single value for the central tendency is the mean (or average) value or the median. Generally speaking, for normally distributed data the arithmetic mean is appropriate, when for non-normal data the medium value usually is the best representative of the central tendency. The worst-case estimate is mostly based on a specific percentile (usually 80th, 90th, or 95th percentile), or on the highest value found in a series of data. Although the choice for a specific percentile is also subjective, the use of a percentile is preferred over the use of an arbitrary high value. In many Risk Assessments, no specific choice for the level of precaution is made; instead, rather arbitrary values are selected on the basis of available data in the literature.

Most outputs from site-specific Risk Assessment, such as calculated human exposure or the number of ecological species affected, must be regarded as indications of truth values. Nevertheless, Risk Assessment is an extremely useful tool, as long as it is smartly used.

First, outputs from Risk Assessments can always be safely used for comparison of risks (*comparative Risk Assessment,* aka: *relative Risk Assessment*), for example, for priority setting. Higher exposure, for example, generally means a higher risk; or to put it even better, a higher *Risk Index* generally means a higher risk. Second, Risk Assessments based on worst-case assumptions can be used in a first step of a Risk Assessment procedure. Generally speaking, this implies that when there is no unacceptable risk, even under these worst-case conditions, it is relatively safe to state that unacceptable risks to human health, the ecosystem, the groundwater or agricultural products are very unlikely. The risk assessor, however, needs to be alert to the fact that the worst-case conditions indeed apply to the specific site. Imagine, for example, that a worst-case Risk Assessment is performed, for the purpose of investigating whether people in a specific residential setting might experience unacceptable risks, by using upper limit estimates for the crucial exposure parameters. If the result from this Risk Assessment under these worst-case conditions shows a Risk Index that does not exceed a value of 1, an unacceptable risk is very unlikely for normal conditions at the site. However, when inhabitants at the residential site, now or in the future, grow a much larger percentage of their vegetables on the site (large gardens), unacceptable risks cannot be excluded.

Another possible pitfall is that the boundary condition 'based on worst-case assumptions' is a subjective criterion, which is difficult to motivate and communicate. The level of conservatism is rarely concretised in Risk Assessments or, at best, at the level of subjective terminology such as 'based on worst-case assumptions'. Moreover, risk assessors sometimes might feel the urge to protect themselves from false negatives (the assumption that there is no unacceptable risk, when in reality there is one) which might lead to an unnecessary over-conservatism.

Scientists and regulators usually are looking for a balance between 'to be sure to be on the safe side', and realism and pragmatism. For this purpose the term 'realistic worst case' is often used, although this still is a subjective criterion. The use of a specific percentile, for example, the 90th percentile of each input parameter representing worst-case conditions, is a more objective criterion. However, the selection of this percentile is also a very subjective process. Choices for specific percentiles (usually 80th, 90th, or 95th percentiles) are often mentioned in Risk Assessments, but are seldom explained.

For more ambitious applications, the risk assessor needs to be aware of the sensitivities and uncertainties that are involved in the Risk Assessment tools. An experienced risk assessor needs to use insight when it comes to the most sensitive input parameters. A *sensitivity analysis* and *uncertainty analyses* can help to systematically identify the most sensitive input parameters, model equations, etc. The risk assessor also needs to be aware of the limitations of the outputs from Risk Assessment and check these against the purpose of the Risk Assessment. When the uncertainties are too great, the performance of additional assessments will be necessary, or the power of the results of the Risk Assessment will have to be adapted to more modest conclusions, that is, by communicating the restrictions and uncertainties.

A relatively simple, though quite time-consuming way of dealing with the lack of reliability, is to follow a *probabilistic* instead of a *deterministic approach*. A deterministic approach, based on point estimates in input parameters and resulting in a single value, does not give any information about the variation in that value. Moreover, since information about the lack of variation is lacking, stakeholders might get a misleading idea about the accuracy involved. In a probabilistic approach, input parameter point estimates are replaced by probability density functions, for at least the most sensitive input parameters. The most popular probability functions are normal, lognormal, cumulative and uniform distributions. Several software packages are available, for example, Crystal Ball, to determine the probability density functions from a series of data.

The procedure for the determination of probability density functions depends on the purpose of the Risk Assessment. For ‘generic Risk Assessment purposes’, for example, the derivation of Soil Quality Standards, the most effective way is to incorporate both uncertainty and variability in the probability density functions. For site-specific applications, however, most of the variability could be eliminated by measurements, so that the probability density functions mainly cover uncertainty. The most popular way of performing a probabilistic Risk Assessment is based on Monte Carlo techniques (e.g., Seuntjes (2004), who assessed the risk of the leaching of cadmium from soil, originating from the former presence of non-ferrous industries, into the groundwater in Lommel, Belgium). Burmaster and Anderson (1994) described 14 principles of good practice to assist people in performing and reviewing probabilistic or Monte Carlo Human Health and Ecological Risk Assessment.

The result of a probabilistic Risk Assessment is a probability density function of an important measure for risks, for example, human exposure or percentage of soil organism affected, or of a Risk Index. The huge advantage of such a probabilistic procedure is that the impact of uncertainties and variability is made transparent in the resulting risk appraisal. However, a choice needs to be made for the level of acceptability, in terms of a specific percentile of the probability density function as output of the Risk Assessment. Although this offers a more sophisticated way of dealing with acceptable risks, there are no objective criteria to underpin this choice.

Since Risk Assessment is a relatively unreliable process, it is of the utmost importance to describe each and every step taken, from the field survey on up to Risk Management solutions. This should be done in such a detailed way that the Risk Assessment is reproducible for third-party risk assessors. The report must explain which political boundary conditions are incorporated in the Risk Assessment. Furthermore, it should refer to all the Risk Assessment tools (including all input parameter values) that were used, along with associated references.

Because of the characteristically limited reliability involved with Risk Assessment, it is recommended to organise, at least for crucial reports, peer reviews and/or second opinions. Peer reviewers cannot eliminate the uncertainties, but they can judge whether risk assessors have made these uncertainties transparent and also, very importantly, whether the uncertainties rectify the conclusions. Several countries include peer review or second opinion procedures in their acts and laws. Alternatively, these procedures are often included in national guidance documents.

1.5.4.3 Validation

The lack of reliability of Risk Assessment results is supported by numerous validation, comparison and round-robin studies. It must be realised, however, that validated models hardly (if at all) exist (Leijnse and Hassanizadeh 1994). In fact, only model applications can be validated. The reason for this is that for each specific model application, different equations and input parameters are the most relevant. Therefore, in each specific model application a different part of the model is tested.

When a smart combination of model applications is validated, however, it at least ensures a level of confidence in the whole model. But since it is a subjective decision as to what kind and how many validated model applications are needed to cover the whole range of possible model applications, and the criterion for ‘validated’ for separate validations is quite vague, the term ‘validated model’ is better off being avoided.

A similar process that investigates the performance of models and procedures is verification. Verification focuses on the testing whether a predefined hypothesis is true.

1.6 Risk Management

1.6.1 Scope

Sometimes, a broad definition of *Risk Management* is followed, that is, Risk Management is the whole risk-based procedure for contaminated site management. According to this broad definition, Risk Assessment is considered as an important component of Risk Management. In this book, however, a more narrow definition of Risk Management is used (see Fig. 1.3) that focuses on the development of the strategies for dealing with the risks, only. From this perspective, the term Risk Management is more directly related to the dictionary definition of management, which includes active words such as ‘handling’ and ‘controlling’, generally with the purpose of bringing contaminated sites back into beneficial use.

Risk Management is appropriate when the conclusion from a Risk Assessment is that a particular risk is unacceptable. It includes avoiding the risks, mitigating or removing risks and, last but not least, communication about the risks with the parties involved. The keyword in Risk Management is *risk reduction*. There are many ways to achieve risk reduction. Basically, Risk Management relates to removal or controlling of the source, that is, *source control treatment*, or to blocking the pathway from source to receptor. The challenge is to find the optimum balance between the most effective and most cost-efficient way of doing this by weighing the short-term advantages against the costs of aftercare.

Remediation (aka: restoration, or clean up), that is in its most strict definition elimination of the source and the resultant soil contamination, is the most direct way of risk reduction. However, remediation often is too drastic an activity, whose results are not in alliance with the social and technical impact at the site and the costs. Alternatively, source control or the application of barriers, that is, a process which eliminates or blocks the source, might be sufficient. In some cases, compliance with policies requires more stringent measures than are absolutely necessary from a risk perspective.

Communicating with all stakeholders is necessary to find the optimal end goal of Risk Management and to define the procedure for how to achieve this. Often, an intensive negotiating process is needed in which decision-makers play an important role.

1.6.2 The Source

The term ‘source’ could do with some further attention. At any contaminated site, a primary source that is responsible for ongoing contamination of the upper soil layer, e.g., a leaking pipeline, an oil spill, waste materials stored on the surface of the soil, must be fully eliminated, when possible. In case of ongoing atmospheric deposition, elimination of the source often is a long-term political process addressing the responsible parties for immissions of contaminants, and it is not always possible. Other sources might be part of agricultural practices, such as the application of fertilizers and pesticides. In that case, an acceptable soil quality would constitute a harsh boundary condition in agricultural soil management.

Contamination of the upper soil layer, in addition to being a potential cause of risk to human beings, the soil ecosystem and to Food Safety, is a source for groundwater contamination. This might lead to the necessity of removal or control of contaminants in the upper soil layer for the purpose of protecting the groundwater.

1.6.3 Procedures

In the late 1970s, Risk Management was often the same thing as complete removal of the contaminants and, hence, of the risks involved. Harsh remediation measures, such as *Dig-and-Dump* (remediation of the upper soil) and *Pump-and-Treat* (remediation of the groundwater) were the most popular mechanisms to achieve this goal. Alternatively, insulation of the contaminants, and hence of the risks involved, was used as a less strict but cheaper solution. Since the early 1990s, the general focus of Risk Management has evolved into the elimination of *unacceptable risks*, which does not necessarily mean complete removal of the contaminants. Today, the remediation objective is often set at a concentration where the risks for human health, the soil ecosystem, the groundwater and/or Food Safety relate to an *acceptable risk level*.

Moreover, the weighing of the end goal of remediation against necessary costs has evolved into the common way of performing Risk Management.

The most simple and generally least expensive solution for contaminated site problems relates to changing the land use, or adapting the layout of the site within the same land use, in terms of blocking the major exposure pathways. An example of change of land use is using cadmium contaminated sites at the border of a municipality for city expansion, which does not allow substantial vegetable production, instead of using it for vegetable gardens or for agricultural purposes. In this way human exposure through vegetable consumption is reduced or eliminated. An example of changing the layout of a site within the same land use is given by a lead contaminated site with a heterogeneous contamination pattern. The human health risks can be substantially reduced when the buildings are situated on the locations with the highest lead contents and the bare surfaces (garden and borders) on the locations with the lowest lead concentrations. In this way exposure of children to lead through soil ingestion is avoided or reduced. A popular option, mainly efficient for immobile contaminants, is covering contaminated hotspots with pavement, grass or any other vegetation, also reducing the possibilities for hand-mouth contact and, hence, exposure through soil ingestion. Another example is found in Arienzo et al. (2004), who revegitated a soil at a former ferrous metallurgical plant in Naples, Italy, for the purpose (among others) of preventing dispersion of metal-contaminated particles by water or wind erosion. Fencing off highly contaminated parts of a contaminated site, as, for example, described in Louekari et al. (2004), for the purpose of avoiding practically any lead exposure near a former lead smelter in Finland, would be a good example of rather drastic measures in regard to adapting the site use.

The disadvantage of changing land use or the layout of the site with the same land use is that concessions often have to be made in regard to the ideal way the site is used.

Moreover, risks for other protection targets should also be investigated. Therefore, this solution often offers limited possibilities.

1.6.4 Remediation Technologies

1.6.4.1 Scope

Remediation is a hard-to-protocollise activity. It is often not feasible to follow a cookbook-type recipe for the design of a remediation plan. The reasons for this are that for every combination of contaminant, site, soil properties and land use, a different remediation technology may be appropriate. Moreover, the execution of one specific remediation technology can be carried out in many different ways. Therefore, the development of the remediation plan typically must be done on a site-by-site basis. Remediation experts often lobby against rigid remediation plans. Instead, they would prefer a remediation approach in which the proceedings develop during the remediation activities.

The basic distinction in remediation technologies is in situ (at the site) and ex situ (off the site) technologies. In situ technologies, mainly applicable to organic contaminants, have the advantage that no transport of soil material is needed. The huge advantage of ex situ technologies is that the physical-chemical treatment of soil is generally more efficient in a factory than on site. In Fig. 1.8, an illustration of an excavation in Bilthoven, the Netherlands is given, as an example of an ex situ remediation.

Bardos et al. (Chapter 20 of this book) give a detailed description of innovative, sustainable remediation technologies.

1.6.4.2 In Situ Remediation Technologies

The US Environmental Protection Agency includes 12 different in situ remediation technologies in their Annual Status report on contaminated sites treatment

Fig. 1.8 An illustration of an excavation in Bilthoven, the Netherlands, as an example of an ex situ remediation (photo: K. Versluijs; reproduced with permission)

technologies (US EPA 2007). These remediation technologies are Bioremediation, Chemical treatment, Electrical separation, Flushing, Multi-phase extraction, Mechanical soil aeration, Neutralization, Phytoremediation, Soil vapour extraction, Solidification/Stabilization, Thermal treatment and Vitrification.

Phytoremediation, using hyper-accumulators to extract contaminants from soil, focuses on metal elimination from the soil (e.g., Vassilev et al. (2004), who gave an overview of the use of plants for the remediation of metal-contaminated soils, including site decontamination (phytoextraction), stabilization techniques (phytostabilisation), and the use of soil amendments to enhance (in case of phytoextraction) or reduce (in case of phytostabilisation) mobilization of metals). Generally, this is a slow remediation technology. An extensive root proliferation increases metal uptake. At too-high metal concentrations, phytotoxicological effects might hamper an efficient uptake. A relatively efficient plant for Phytoremediation is Brassicaceae, which has a high metal uptake affinity and a relatively high tolerance to metals. Brassicaceae was used, for example, by Kidd and Monterroso (2005) for the purpose of extracting metals from mine-soil material in Spain. Robinson et al. (2000) demonstrated the possibilities of willow (Tangoio) and poplar (Beaupré) clones for phytoremediation of cadmium-contaminated sites.

Electrical separation is based on an electric field in the soil between inserted electrodes, which forces the migration of pore water or groundwater, including metals and organic contaminants. It is a relatively new technology and, therefore, still in the experimental stage, for which further development is necessary. It might, however, be an alternative for the remediation of clayey soils when 'Pump-and-Treat' methodologies are not efficient. Amrate et al. (2005), for example, demonstrated a successful migration of lead in a highly contaminated soil near a battery plant in Algiers, Algeria, where EDTA was added to enhance lead transport.

A spectacular extensive remediation technology that has gained enormous popularity since the mid 1990s is based on biodegradation of organic contaminants and dilution, and is often called *Natural Attenuation*. Indigenous or cultured organisms can be used for biodegradation. In spite of the sometimes high starting costs, the overall budget for this Risk Management procedure is generally low. In addition, it allows for a minimal disturbance of the natural conditions in the soil or groundwater, and there are limited engineering activities needed at the site. Moreover, it even offers opportunities for difficult sites with clayey soils and difficult contaminants such as chlorinated hydrocarbons, even under anaerobe conditions. Natural Attenuation is often combined with ex situ remediation techniques, such as removal of the source. The adage is: use the natural self-cleaning capacities of the soil as much as possible, stimulate natural conditions when necessary and use ex situ remediation technologies only when strictly needed. The success of Natural Attenuation depends primarily on the type of organic contaminant and the performance of the soil ecosystem. The latter depends on the organisms present. Zytner et al. (2006), for example, demonstrated the important contribution of fungal metabolism for the degradation of branched hydrocarbons. For this reason the chemical and physical characteristics of the soil and the artificial oxygen and nutrient supply are dominant factors.

In situ Bioremediation of organic contaminants is especially difficult in low permeability soils. Athmer (2004) described a procedure for integrating electro kinetics with in situ treatment for the remediation of TCE (trichloroethylene) contaminated clay soils in Paducah, Kentucky, USA, to address this problem. It generated a uniform migration of trichloroethylene through the soil to treatment zones.

Peter et al. (Chapter 22 of this book) give a detailed description of Natural Attenuation and of its practical possibilities.

As specific applications, 'bio-screens' are used, that is, zones with an active, often stimulated, degradation at strategic positions in the soil system, or Funnel-and-Gates techniques, in which contaminants are led to zones with an active degradation.

Several materials have been proven to be effective in *Solidification/Stabilization* (aka: *immobilization*, or fixation) of heavy metals in soils. A proven method to fixate metals in soils is mixing the soil with lime (liming) or cement. Yukselen and Alpaslan (2001), for example, successfully immobilized copper, and iron in soils in an old mining and smelting area located along the Mediterranean coast in northern Cyprus. They showed that an additive/soil ratio of 1/15 (on mass basis) resulted in the optimal immobilization, for both lime and cement. This ratio very much depends, of course, on the soil type and soil properties, mainly pH. Tlustoš et al.

(2006) demonstrated that by the addition of lime and limestone to a contaminated Cambisol with 7 mg/kg cadmium, 2,174 mg/kg lead and 270 mg/kg zinc, the mobile (0.01 mol/l $CaCl_2$ extractable) fractions dropped by 50, 20 and 80% for cadmium, lead and zinc, respectively. The pH was increased from 5.7 to 7.3. Consequently, the metal concentrations in straw and grains of wheat were significantly reduced.

Száková et al. (2007) found substantial differences in reduction of the mobile (0.01 mol/L aqueous $CaCl_2$) fractions of metals when applying lime, limestone, and zeolite to contaminated soils. However, although the mobile fraction of cadmium and zinc indeed decreased, the mobile fraction of lead was hardly affected and the mobile fraction of arsenic even increased in some of the treated spots. The availability of arsenic was more affected by different characteristics of experimental soils than by individual soil amendments. Moon et al. (2004) contributed the fixation of arsenic to inclusion of arsenic in pozzolani cement reaction products and the formation of calcium-arsenic precipitates.

Ameliorating soil materials can be of natural origin, such as clay or bauxite residue. Alternatively, several by-products of production processes are used for this purpose. Red mud, a by-product of the aluminium industry, for example, has been identified as an effective amendment for in situ fixation of heavy metals in soil because of the high content of Fe and Al oxides (Zhang et al. 2002). Friesl et al. (2003) demonstrated the efficiency of the amendment of red mud (10 g/kg) in four soils, in the vicinity of a former Pb-Zn smelter in Austria, highly polluted with (among others) Zn (2,713 mg/kg) and Cd (19.7 mg/kg). This resulted in the reduction of metal extractability of 70% for Cd and 89% for Zn.

Other cements used as fixation material are sulfoaluminate cement, powdered activated carbon, quick lime (Guha et al. 2006), and ferrous sulphate (Warren et al. 2003). These authors showed that accelerated carbonated treatment substantially reduces the availability and, hence, the risks, of mercury in soil.

One disadvantage of the amendment of immobilizing soil materials may be the presence of other contaminants, which implies that the immobilization of specific metals is accompanied by the introduction of other contaminants. Red mud, for example, includes arsenic, chromium and vanadium (Friesl et al. 2003). Therefore, the optimal application of immobilizing materials to soils requires the optimum balance between an effective binding of metals and minimizing the negative effects of other contaminants. Friesl et al. (2004), for example showed that, at a red mud addition of more than 5% of total soil weight, the disadvantages of introducing other contaminants exceeds the advantage of fixation of metals.

Grotenhuis and Rijnaarts (Chapter 21 of this book) give a detailed description of in situ remediation technologies.

1.6.4.3 Ex Situ Remediation Technologies

The US Environmental Protection Agency includes 14 different ex situ remediation technologies in their Annual Status report on contaminated sites treatment technologies (US EPA 2007); these are Bioremediation, Chemical treatment, Incineration, Mechanical soil aeration, Neutralisation, Open burn/Open detonation, Physical separation, Phytoremediation, Soil vapour extraction, Soil washing, Solification/Stabilization, Solvent extraction, Thermal desorption, and Vitrification.

1.6.4.4 Barriers

Barriers are used to isolate contaminants at contaminated sites from the surroundings and, hence, to protect any protection targets in the surroundings of the contaminated site. Compacted soil materials are recognized liners. Clay, or clayey soil material, is the most obvious natural barrier material. They both have a low hydraulic permeability and a chemical buffering capacity through adsorption. Kabir and Taha (2004), for example, demonstrated an effective barrier function of compacted sedimentary granite residual soil material for the isolation of contaminants in landfills. They showed that this material has a hydraulic conductivity lower than the suggested limit (1×10^{-7} cm/s) of the various waste regulatory agencies in the USA. In addition, it has adequate strength for stability, and exhibits small shrinkage potential upon drying. Qian et al. (2002) specified the requirements of soil materials as effective barriers, in terms of contribution of silt and clay, plasticity, and limitations to the contribution of gravel-size materials and chunks of rock.

Several waste materials such as fly ash could be used a barrier material. Sivapullaiah and Lakshmikantha (2004), for example, demonstrated that the addition of bentonite to fly ash improves the chemical buffering function and the geotechnical properties of the barrier. Kaolonite and bentonite (a commercially available high swelling clay) are artificial alternatives for barrier materials.

1.6.5 Ecological Recovery

A problem with many remediations, especially ex situ remediation measures, is that the soil ecosystem, vegetation and above-ground fauna generally are negatively impacted, at least on the short term. In case of Dig-and-Dump technologies, the habitats, the organisms and the seed pool are removed from the site. Ex situ thermical treatment results in total elimination of soil organisms and organic matter. Especially when a remediation is triggered because of unacceptable ecological risks, the question is relevant if the ecological benefits on the longer term counteract the negative impact on the short term. The possibilities and timeframe for ecological recovery strongly depend on the type of soil that is applied in the final stage of the remediation. Especially clayey soil material high in organic matter speeds up the recovery process. The application of comparable soil as the wider environment, however, improves the development of a regionally appropriate ecosystem. The effects on above-ground fauna can be reduced through a stepwise remediation procedure, in which in different stages only a part of the site is remediated, so that recolonisation of organisms in the 'new soil material' can take place from the parts that have not yet been remediated.

1.6.6 Remediation Objectives

An end goal of a Risk Management procedure must be defined, and generally expressed as soil or groundwater concentration. Although the term *Risk Management objective* would have been a more appropriate term, this intended soil or groundwater concentration generally is called *remediation objective* (aka: *remediation goal*).

In the early days of contaminated site management, in the late 1970s, the remediation objectives were commonly set at the zero level, often not supported by any explicit considerations. Today, the goal of Risk Management generally relates to an *acceptable risk level* for the relevant protection targets. The selection of protection targets and the definition of an acceptable risk level are beyond the scope of science and are the responsibility of decision-makers. Contrary to the appraisal of existing soil contamination, which relates to imposed risks, Risk Management is supposed to 'create' a desirable situation, versus certain efforts and costs. Therefore, there are good reasons to select more protection targets and more stringent protection levels for the objectives of Risk Management than for curative decisions on existing soil quality.

Examples of *acceptable risk levels* are the Negligible Risk (NR) for human health as a target for the soil upper layer, or the Negligible Risk (NR) for the aquatic ecosystem as a target for the groundwater. Another option for a remediation target that is not based on risks is the (natural) background concentration (as a target for the upper soil or the groundwater), or commercial production criteria as a target for agricultural products.

The process of deriving remediation objectives includes the following steps:

- Selection of protection targets.
- Definition of 'policy requirements' for each protection target (e.g., 'it must be possible to grow the complete vegetable package of a family in a vegetable garden'; or 'the soil ecosystem must be fully protected in a nature reserve').
- Translating the 'policy requirements' into Risk Assessment terms (e.g., in analogy with the examples above, 'exposure through the complete vegetable consumption from the own garden equals the Reference Dose for exposure'; or: '95% of the soil ecosystem must be protected (in that case, an affected fraction of 5% is assumed as "full protection")'.
- Derivation of risk limits in soil or groundwater for every protection target, and for all selected contaminants.
- Selection of the appropriate risk limit in soil or groundwater as a remediation objective (usually the lowest of all risk limits in soil or groundwater) for all selected contaminants.

Except for soil concentrations, alternative types of remediation objectives could be defined. Von Lindern et al. (2003a), for example, used the lead concentration in house dust as a goal for the remediation of the Bunker Hill Superfund site in northern Idaho, USA, since exposure through dust ingestion has been recognized as a principal exposure pathway. Von Lindern et al. (2003b) focused on lead blood levels as a remediation objective of the Bunker Hill Superfund site.

In some cases there are good reasons for focusing on lower soil concentration levels as remediation objectives than is strictly needed for human health protection. Several remediation technologies, such as Dig-and-Dump for example, do not always allow for 'gradual' risk levels after remediation, but may result in a clean soil.

1.7 A Closer Look into Risk Assessment

1.7.1 Types of Risk Assessment

1.7.1.1 Purpose

Generally speaking, contaminated site Risk Assessment offers two possibilities. First, Risk Assessment can be used to investigate a specific site. This type of Risk Assessment is called *site-specific Risk Assessment* or *actual Risk Assessment*. In this case, information about the specific site is available. Second, Risk Assessment can be used to derive Soil Quality Standards. This type of Risk Assessment is often called *potential Risk Assessment* or *generic Risk Assessment*. Often, generic Risk Assessment is the first step in Risk Assessment frameworks, followed by site-specific Risk Assessment when the generic Risk Assessment does not result in a clear decision as to risks.

1.7.1.2 Site-Specific Risk Assessment

Risk Assessment related to a specific site is often called *actual Risk Assessment*. From this perspective, 'actual' is used in the sense of 'existing in fact' and not necessarily in the sense of 'existing at this moment'. It is possible, for example, that a Risk Assessment might be performed for the purpose of investigating whether it is 'safe' to reside at a specific residential site, which might be contaminated. In that case, the relevant time frame for Risk Assessment can vary from several years up to several decades, depending on the time frame over which the specific contaminants reveal effects. Therefore, it does not make sense, for example, to focus on the actual layout of the garden, with or without vegetables grown for one's own consumption, since this layout may change over a period of years or decades. In fact, an assumption needs to be made for a representative contribution of vegetables from one's own garden, independent of the situation at the time that the Risk Assessment is performed. Specifically, in situations in which a Risk Assessment is performed for the purpose of investigating the risks for a future land use, it does not make sense to base the Risk Assessment on features that relate to the present land use. For this reason, the term 'site-specific Risk Assessment' is used in this book, rather than 'actual Risk Assessment'.

Nevertheless, for site-specific Risk Assessment, relevant information might be available that facilitates the assessments of risks for human health, the soil ecosystem, the groundwater, or Food Safety. Although not all information present might be relevant, a huge asset is that measurements can be performed, which significantly improve the quality of the Risk Assessment, and reduce uncertainties (see Section 1.7.3).

1.7.1.3 Potential Risk Assessment

The term ‘potential’ is not to be missed in any list of vague terminology. Paradoxically, it also is a very useful term in contaminated site management. Literally, in the sense of ‘possible, when certain conditions apply’, there is a potential unacceptable risk at any site where contaminants have been measured by definition, independent of the concentration. The ‘certain conditions that apply’ are, for example, an intensive contact of human beings with the soil with regard to human health risks, a relatively high bioavailability of the contaminants in soil with regard to Ecological Risk Assessment and Food Safety, or a low pH and the presence of metals with regard to Groundwater-related Risk Assessment.

A potential unacceptable risk also might refer to the fact that a site-specific Risk Assessment has resulted in the conclusion ‘unacceptable risk, with a low level of reliability’. This could be the case, for example, when only a first step in a wider Risk Assessment framework has been performed, with conclusions based on the chemical analyses reports of a limited number of samples. However, since risk-based soil quality assessment is characterised by substantial uncertainties in general, the adjective ‘potential’ in the meaning of ‘conditional’ could practically always be added. The only benefits of this use of the term ‘potential’ would be to stress the lack of reliability to the stakeholders. This function increases the level of confusion rather than supporting the Risk Assessment and the Risk Communication. Therefore, ‘potential’ should not be used for the purpose of alerting those involved to the limited reliability of a Risk Assessment.

A more appealing use of the phrase *potential Risk Assessment*, in this context also referred to as *generic* Risk Assessment, is related to the derivation and use of Soil Quality Standards. Since Soil Quality Standards are not focused on a specific site, but rather relate to a whole series of unknown contaminated sites, these Soil Quality Standards must be derived from generic scenarios.

1.7.2 Soil Quality Standards

Soil Quality Standards (aka: (soil) Guideline Values, (soil) Screening Values, or Target Levels) are generic values enabling a distinction into two classes for which the measured concentrations in soil are either higher or lower than the Soil Quality Standard. They can be considered as the core of contaminated site management. In the early days of contaminated site management, a list with Soil Quality Standards was about the only appraisal framework available, and often used for the separation between acceptable and unacceptable cases of soil contamination. Since the late 1980s, *risk-based* Soil Quality Standards have been derived in several developed countries.

Several types of Soil Quality Standard exist, for different purposes. Carlon and Swartjes (2007b) distinguished three classes of Soil Quality Standards. The first class of Soil Quality Standards, with the most stringent values, represents the upper limit for long-term sustainable soil quality, appropriate for prevention purposes or as remediation objectives. The second class, with the highest values, triggers actions such as either a more detailed Risk Assessment or Risk Management actions (e.g., remediation) when exceeded, and are used for curative purposes, that is, for supporting the risk appraisal for existing contaminated sites. The third class is an intermediate class, and supports further research actions such as the performance of more detailed soil sampling.

As was mentioned in Section 1.7.1.3, Soil Quality Standards are applied to a whole series of contaminated sites. Therefore, a generic exposure scenario needs to be defined for a hypothetical site. Generally speaking, such a generic scenario either relates to standard assumptions, as for frequently found contaminated sites, or to conservative assumptions. The latter must certainly be the case when ‘false negatives’ (the incorrect assumption that there is no unacceptable risk) get a higher political negative weight than ‘false positives’ (the incorrect assumption that there is an unacceptable risk). Also in a case where Soil Quality Standards are used as a trigger for possible site-specific Risk Assessments, generic scenarios as basis for the Soil Quality Standards need to be based on conservative assumptions.

A variation on generic Soil Quality Standards relates to ‘land use-specific Soil Quality Standards’. As the term says, it refers to several Soil Quality Standards for different land uses, for each specific contaminant. One advantage of human health-based land use-specific Soil Quality Standards is that more realistic exposure scenarios for the respective land uses can be used. An advantage of ecologically based land use-specific Soil Quality Standards is that a more appropriate level of ecological protection can be chosen for the respective land uses. The disadvantage of land use-specific Soil Quality Standards is that the derivation process is much more intensive, since a series of Soil Quality Standards must be derived for each contaminant. The use of land use-specific values in practice is less convenient, since a choice needs to be made for each site as to which land use is appropriate. Moreover, the application of land use-specific Soil Quality Standards may give a misleading idea of accuracy.

Examples of human health-based Soil Quality Standards are given in Hristov et al. (2005) for *Human Health Soil Screening Levels (CHHSSL)* in California, USA, and in DEFRA and EA (2002) for *Soil Guideline Values* for metals in the UK. Examples of ecologically based Soil Quality Standards are given in Canadian Council of Ministers of the Environment (1999) for Canada, and in National Environmental Protection Council (2003) for Australia. An example of combined (human health and ecological based) Soil Quality Standards is given in Ministry of VROM (2008) for the Netherlands.

Worldwide the Soil Quality Standards used differ to a large extent. This is partly due to differences in the purpose of the Soil Quality Standards, but the technical frameworks also show many differences. Provoost et al. (2006) compared Soil Quality Standards for eight metals and metalloids, from Canada, Flanders (Belgium), France, Germany, the UK, the Netherlands, Norway, Sweden, Switzerland and the USA. For most contaminants they found differences between the highest and lowest value of more than a factor of 1000. They concluded that some of these differences could be explained by political differences, such as the choice of protection targets and risk levels. Some of the other differences between Soil Quality Standards, however, are explained by technical/scientific differences between the procedures used in the different countries. Swartjes and Carlon (2007) came to similar conclusions for Soil Quality Standards used in 16 European countries. They found even higher differences for organic contaminants. Swartjes (2007) concluded that differences between seven European exposure models, important instruments in the derivation of human health-based Soil Quality Standards, can result in widely different risk appraisals for the same exposure scenarios, especially for contaminants that are mobile and even more for contaminants that are volatile. Therefore, the need for a higher consistency of Risk Assessment tools is acknowledged in Europe (Swartjes et al. 2009).

1.7.3 Measurements

Measurements in contact media can significantly improve the quality of a site-specific Risk Assessment. The kind of measurements that are possible and the benefits of these measurements vary. In this section, a general view concerning measurements in contact media will be given. In the introductionary chapters on Human Health Risk Assessment (see Chapter 6 by Swartjes and Cornelis, this book), Ecological Risk Assessment (see Chapter 13 by Swartjes et al., this book) and Groundwater-related Risk Assessment (see Chapter 17 by Swartjes and Grima, this book) the most important measurements are described in more detail.

First, *direct input parameters* could be measured, such as the concentration in human blood or body tissue (Human health Risk Assessment), or the number of earthworms in soil (Ecological Risk Assessment). Second, *basic input parameters* could be measured, such as the concentrations in soil compartments, or in contact media. Third, s*upportive input parameters* could be measured, such as soil properties or input parameters that relate to long-time human behaviour, if relevant for the site over the relevant time span.

Measuring of supportive, basic or direct input parameters have both advantages and disadvantages. Generally speaking, the reliability of the assessment of the *actual* Risk Assessment improves when more supportive measurements are available, even more so when basic input parameters are measured and the most often when direct measurements are performed. The disadvantage, however, is that measurements are expensive and generally more expensive in the order of supportive, basic and direct input parameters, and, therefore, not always suited for routine Risk Assessment.

Moreover, the general belief that measuring gives a more reliable value than a calculation is often true, but not always. In some cases, measuring a representative value is extremely difficult for technical reasons and due to spatial and/or temporal variation. Moreover, measurements are strongly related to the *present* situation and may not represent the long-term conditions for which it is stated, for example, that 'the site is safe to live on'. The risk assessor needs to find a smart balance between the determination of a more reliable measured input parameter that might be less representative for the long-term risk, or a less reliable calculated input parameter that does represent long-term site conditions. Obviously, the risk assessor also has to account for the additional costs of measurements when formulating the involvement of measurements.

An example of such a dilemma is the Human Health Risk Assessment for a well-maintained (manuring, liming) vegetable garden. The risk assessor has to decide if the Risk Assessment benefits from measured concentrations in the vegetables present at the site, which are relatively accurate for the present situation, but may give underestimations for future situations in which another owner neglects soil liming. Alternatively, the risk assessor could perform relatively unreliable calculations of the concentrations of a representative combination of vegetables of choice, on the basis of total soil concentrations and soil properties belonging to an appropriate, that is, average or neglected liming conditions.

Generally speaking, measurements of direct input parameters are most valuable for Risk Assessments that relate to, let's say, the first few years (maybe one to three years), while measurements of basic input parameters are useful for the subsequent few years, see Fig. 1.9. As illustrated in this figure, supportive input parameters often represent time spans up to decades.

This picture only gives a general insight into the time spans for which measurements are useful, mainly for the purpose of illustrating the importance of the time frame when deciding on measurements. Direct measurements for Human Health Risk Assessment, such as measurements of the cadmium levels in blood, are often

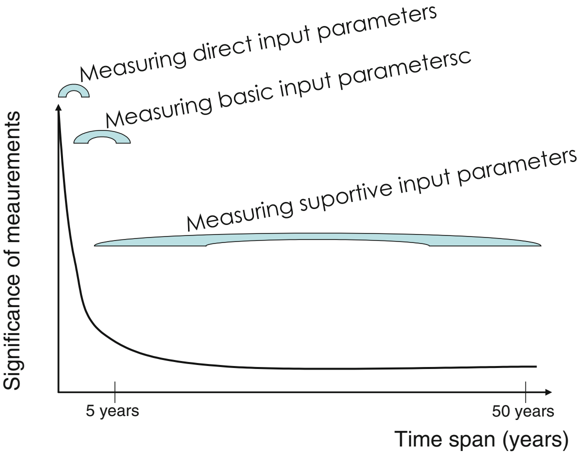


Fig. 1.9 Significance of measurements as a function of the time span for which the Risk Assessment applies (*straight line*). Indication of the time span, for which different type of measurements are useful (*dark shaded arches*)

useful for short-term risk appraisal. Measurements of cadmium in urine, however, represent a measure of cumulative lifelong exposure. Since this represents the site exposure history, is it defensible to assume this measure also is representative for long-term future exposure, when no changes in exposure conditions are expected. And with regard to supportive measurements, the representative time span of clay content (decades to centuries) and pH (1–10 year), for example, largely differ.

1.7.4 Laboratory Data Versus Field Data

Appropriate Risk Assessment is all about what is happening in the real world, that is, at contaminated sites. However, Risk Assessment would hardly be possible without the support of laboratory experiments. These experiments are used for at least three different purposes. First, for ethical and technical reasons, direct human toxicological effect data are rarely available. These data need to be derived from laboratory experiments with animals, following strict guidelines with regard to laboratory animal welfare. Also ecological effect data heavily depend on laboratory experiments, since it is difficult to investigate effects on one specific species in the field. Second, it is inconvenient to control standard environmental conditions in the field.

As a consequence, Human Health and Ecological Risk Assessment both strongly depend on experimental laboratory data. To transfer the laboratory effect data to human health, and ecological effect data and to real world applications (field conditions), the laboratory effect data are divided by *assessment factors* (often called extrapolation factors). These assessment factors, which are often in the range of 10 to values as high as 10,000, cover intraspecies and interspecies differences and, last but not least, uncertainties associated with laboratory studies, which always cover a limited amount of tests, for a limited time span. A specific case of intraspecies differences in Human Health Risk Assessment is the differing sensitivity within the human population. In this case, an assessment factor could be applied, when politically feasible, to also protect the most sensitive part of the human population. In case of effects data for a whole ecosystem, instead of a single species, uncertainties due to interspecies variation are dominant. In fact, a major problem in Ecological Risk Assessment is the extrapolation of observations from individual and population levels to the ecosystem level (Eijsackers et al. 2008). Again, assessment factors can be used to cover the corresponding uncertainties.

A second important use of laboratory studies relates to the assessment of input parameters. These input parameters may include ‘supportive parameters’ such as physico-chemical contaminant characteristics (vapour pressure, octanol-water partition coefficient (Kow), water-saturated permeation coefficients, etc.). But laboratory measurements can also focus on ‘basic’ (more lumped) input parameters, such as indoor air concentrations, or leachate concentrations. Often, such measurements are prescribed in manuals, guidance documents or Decision Support Systems. Examples of this are the measurement of the concentration in vegetables in Tier 3 of the Dutch tiered approach used to determine the risk due to the vegetable consumption

(Swartjes et al. 2007), or the performance of bioassays in several tiers of the TRIAD approach, to determine the site specific ecological risks (see Chapter 15 by Rutgers and Jensen, this book).

Finally, a third purpose of laboratory studies is validation of models or testing of technologies (pilot studies). An example of a validation study is the comparison of calculated with measured indoor air concentrations for 11 petroleum hydrocarbons and chlorinated solvents sites in the USA and Canada (Hers et al. 2003).

In short, to enable Human Health and Ecological Risk Assessment, experiments are simply a necessity. In principal, laboratory experiments for validating and testing and, to a lesser extent, for assessing input parameters could be replaced by field studies. However, the adage that ‘field studies are always better than laboratory studies’ requires a few nuances. The choice for field versus laboratory studies simply depends on the trade-off between the control of conditions in the laboratory versus the degree of reality in the field. Of course, financial arguments also must be included in this choice. Often a multitude of experiments can be performed in the laboratory for the same budget as for one single field study. Another, non-scientific, aspect that is in favour of field experiments is that they better connect with public perception, but also even sometimes with scientific perception.

1.7.5 Expert Judgement

It is a major inconvenience when in a scientific discipline the possibilities for a quantitative analysis are lacking. In such a case, *expert judgement* may offer an alternative. Expert judgement is the process in which experts determine an opinion, (partly) based on ‘gut feeling’. In an optimal expert-judgment process, several experts are involved, and opinions are based on consensus.

A situation, in which expert judgement can be used, for example, is when a Soil Quality Standard for a specific contaminant is lacking and effect data are not available for this contaminant. Via expert judgement a ‘substitute contaminant’ with similar physico-chemical contaminant characteristics can be selected, for which it is expected that effects, and hence the Soil Quality Standard, are in the same order of magnitude. Another example relates to the optimal balance between the use of a few available measured contaminant concentrations for a specific vegetable and of many less appropriate data for a non-edible crop, in terms of determining a representative concentration for a specific vegetable.

In practice, expert judgement can vary from the opinion forming of a single expert up to striving towards consensus within a group of appropriate experts. A decent expert judgement must be performed within a well conceived group, that often includes experts from different disciplines and, preferably, individuals who approach the issue from a different angle. In any case, the risk assessor needs to describe the expert judgement process and the decisions criteria that were used so that scientists, regulators, and other stakeholders can decide how to weigh the outcome of the expert judgement.

1.7.6 Essential Metals

Several contaminants that are found in soil, mainly metals, are essential for the functioning of human beings or organisms. Twenty-five naturally occurring elements are believed to have an essential function in plants and animals. As a consequence, humans and/or organisms need to take these contaminants up, for example in their diet. Among these *essential metals* are zinc, copper and selenium. It is interesting to note that about 2 billion people worldwide, mainly children and women, suffer from zinc deficiency, partly due to low zinc concentrations in soils in Africa, Latin America and in some parts of Asia (Prasad 1998; Ramakrishnan 2002). Most organisms have a narrow range between optimal concentrations (toxicity) and deficiency.

From the perspective of essentiality, it seems a paradox to name essential metals 'contaminants'. However, analogous to the definition used in this book in Section 1.3.1, that is, that the toxicity of chemicals depends on the dose at which humans or organisms are exposed, the period that this exposure takes place and the frequency of exposure and the form (speciation) in which the chemical substance is available, these essential contaminants are not beneficial by definition. Generally speaking, however, essential contaminants are beneficial at specific low doses.

There has been a lot of debate about Risk Assessments in which essential metals are involved. However, at high concentrations and, hence, toxic levels, risk decisions must not be influenced by the fact that the same contaminants would have been useful at lower concentrations. Of course, in the definition of an end goal of Risk Management (e.g., remediation objectives), the essentiality of contaminants must play an important role.

1.7.7 Background Concentrations

Many definitions are used for *background concentrations* (aka: baseline values or Reference Values), often in combination with the adjective 'natural' (natural background concentrations). Generally speaking, background concentration refers to the concentration in soil or aquifer over a larger area (that is, on a larger scale than the site that is under investigation). These background concentrations can either be related to natural processes or to anthropogenic activities (that is, anthropogenic activities, which took place during years, decades, or even centuries, other than the activities that caused the contamination under investigation). In the context of this book, the distinction is followed as described by the US Environmental Protection Agency (US EPA 2002):

- Natural background concentrations: contaminants present in the environment in forms that have not been influenced by human activity.
- Anthropogenic background concentrations: human-made or natural contaminants, present in the environment as a result of human activities (but not specifically related to the site in question).

Anthropogenic background concentrations can relate to both man-made and natural contaminants. The reason for this is that human activities can be responsible for the release of man-made contaminants over larger areas, for example by large scale exhaust of man-made organic contaminants via the air, followed by atmospheric deposition. But, human activities also can also lead to the spread of natural contaminants in soils, for example, by the soil surface rising, in the case of deep polder areas with shallow groundwater tables, exposing soil material that has been naturally contaminated. So the distinction is based on the type of activity that caused the background concentrations and not on the origin of the contaminants.

Like every classification of background concentrations, this distinction is artificial, since there is a transition from one to the other, for example, human manipulation of the flow characteristics of a river that will impact the deposition of sediment with natural (or maybe also anthropogenic) contaminants. Moreover, some contaminants may contribute to the background concentration as a result of both natural processes and man-made activities, such as the combined presence of naturally occurring arsenic and arsenic from pesticide applications or smelting operations.

Naturally occurring contaminants are often metals. Most organic contaminants are man-made, although many exceptions are known such as PAHs (polycyclic aromatic hydrocarbons) that are formed through natural processes such as wood fires. Cyanide-containing chemicals are produced by a wide range of organisms and plants as part of their normal metabolism. Bacteria and fungi are known producers of cyanide. A few species of centipedes, millipedes, insects, beetles, moths and butterflies secrete cyanide for defensive purposes in repelling predators such as toads and birds (MERG 2001). Some of the common plants that contain cyanide are cassava, sweet potatoes, corn, lima beans, almonds, radishes, cabbage, kale, brussels sprouts, cauliflower, broccoli, turnips, lettuce, kidney beans, and it can be found in the pits or seeds of cherries, plums, apricots, pears and apples (MERG 2001).

Many statistical procedures exist to determine the background concentrations. In general, the type of contaminant (natural or anthropogenic) does not influence the statistical or technical method used to characterize background concentrations (US EPA 2002). In most countries, background concentrations have been established on a national or regional level (e.g., Lavado et al. (2004), who determined metals background concentrations in Pampas soils in Argentina).

Background concentrations can, at the least, be used for two different purposes. First, in relation to Ecological Risk Assessment, it is often assumed that the background concentration does not pose any risk, or less risk, to the soil ecosystem, since the organisms are adapted to the long-term situation (‘Added Risk Approach’). Therefore, Risk assessors need to exclude or nuance the risks caused by the background concentration. Second, background concentrations can be politically used for the definition of ‘acceptable’ (from a political perspective, not in term of risks) soil quality, for example, as an end goal of remediation. The substantiation for this is the fact that large scale remediation of areas of several square kilometres, for example, is nearly impossible for practical and financial reasons. Besides, it is not always defensible, from the standpoint of fairness, to upgrade the soil quality in one small part of the region (the contaminated site), while the surroundings of the site remain slightly contaminated. Therefore, the background concentrations are often declared to be of (politically) acceptable soil quality, independent of the risks involved.

1.7.8 Spatial Scale

It is important to realise on what scale soil contamination needs to be considered. When a backyard is contaminated, it is clear that the human health of the inhabitants needs to be protected. In fact, for every contaminated site, the health of humans in contact with the site must always be the primary focus. In many cases, ecological protection will also be of concern. It is generally wise, certainly for the definition of remediation objectives, to take the soil quality of the wider environment into consideration. It often happens, mostly in urban areas, that remediation of small contaminated sites results in a number of ‘clean’ sites in an otherwise (slightly) contaminated area. This raises the question of whether human health has been sufficiently protected in the slightly contaminated areas around the ‘clean sites’, since the experts have found it necessary to reduce the risk at this specific sites to lower levels than the individuals at the slightly contaminated sites are experiencing. The same issues play a role in the case of a remediation that has taken place for protection of the soil ecosystem.

Moreover, debate could arise about the cost-efficiency of several small-scale remediations in the same area. A regional-scale approach offers possibilities from a cost-efficiency perspective. It is often much less expensive, for example, to investigate and perform Risk Management options for a larger area, in one big project, than to do this in several smaller projects, at different moments in time. And in the case of groundwater contamination, as another example, it is efficient to investigate and manage all the sites that drain towards the same groundwater body and not just to focus on management of the individual contaminated plumes. Moreover, since contaminated groundwater migrates, groundwater plumes have often intermixed, which technically would not make a site-specific approach possible.

Another example in which regional-scale thinking is beneficial relates to *soil material transposition*. Soil material transposition from an intensively used residential site to a less intensively used business park in the same area, for example, would imply risk reduction for the whole area. Obviously, Risk Management is much more complicated than this simple example shows, since more elements are involved than just human health risks, for example, the effects of leaching into the groundwater at the residential site and the business park and juridical aspects in case the locations are situated in different municipalities. However, the example shows that a more regional approach offers practical possibilities for efficient contaminated site management.

A specific example of regional Risk Assessment and Risk Management is the dredging of sediment materials. When dredged materials are deposited on the site of the water courses, the overall contaminant load stays the same within the area. However, since the physico-chemical environment of the water-saturated sediments is completely different than in the water-unsaturated upper soil layer, there are often (but not always) possibilities for risk reduction.

It must be noted, however, that the interest of the stakeholders, as well as the public concern, diminishes when the area under investigation is approached from the perspective of a larger scale. For obvious reasons, the stakeholders and the general public are more committed to the site they own, live on or work on, than the wider area around the site. The consequence for contaminated site management at the scale of the site and the surroundings, and even more so on a regional scale, is that a policy becomes necessary in order to force stakeholders to follow a 'large scale thinking' approach.

1.7.9 Time Domain

In Section 1.1.4.2, it was explained that in the phase of the Problem definition it is very important to define the time frame for which the conclusions from the Risk Assessment applies, since factors that impact human health risks and ecological risks will change over time. Outcomes of the Risk Assessment should usually represent the risks over longer periods. When a site receives a positive risk appraisal, it often is assumed that the site is suited for its purpose for many decades. As a consequence, assumptions need to be made for factors that change over time, mainly with regard to the layout of the site, human behaviour characteristics and the bioavailability of the contaminants.

Moreover, the effects on humans and on bigger animals might only reveal themselves years or decades after exposure. Therefore, it is essential to focus on a toxicologically relevant time frame in Risk Assessment and Risk Management.

With regard to ecological Risk Management the time frame for which ecological restoration takes place is important. For Groundwater-related Risk Assessment, it must be realised that transport times of contaminants might take decades or even centuries. It is essential to be aware of these long time frames, both politically (what time frame is relevant?) and technically.

In fact, changes in the contaminant concentrations due to migration and degradation in the soil and groundwater also need to be considered in Risk Assessments. Remarkably, this is not often done; the concentrations measured are often considered constant in time. Since the concentrations in soil generally decrease over time due to leaching, volatilisation, and degradation, this can be considered as a worst-case approach. This is not always true for groundwater, however, since leaching from soil could increase contaminant concentrations in groundwater.

Many of the Risk Assessment factors also need to be adapted for the determination of exposure scenarios or risk estimates in case the Risk Assessment relates to a different future land use or a different layout of the site under the same land use. This situation frequently occurs, especially in densely populated areas such as in Northwest Europe, since land use transitions are common and are often preceded by a site investigation and, hence, a Risk Assessment.

1.7.10 Costs of Soil Contamination

Large amounts of money are involved in the investigation, and certainly remediation, of contaminated sites. Many developed countries spend amounts in the nine or ten digits (in euros or dollars) on contaminated sites. Today, development firms are used to incorporating cost for soil contamination in their overall costs estimates.

The general rule in most countries that have regulations concerning contaminated site management is that the persons or parties that cause the soil contamination today are fully responsible for the Risk Management solutions, often including recovery of the site to its original situation, that is, eliminating all contamination. This principle is called the *precautionary principle*, see Raffensberger and Tickner (1999) for an overview. Generally speaking, they are obligated to carry all the costs involved.

The cost aspect is much more complex for *historical soil contamination*, that is, contaminations that were caused before any regulations were available. In most countries, the *polluter-pays principle* is the foremost one. This means that the company, association, government or individual that caused a soil contamination is responsible for the costs of investigation and Risk Management. In many cases, however, it is difficult to identify the responsible party. In other cases, the polluter apparently responsible is known, but it is difficult to prove that they indeed caused the soil contamination. This has led to court cases with regard to the identification of the polluter responsible, often between a governmental body and a company.

From the perspective of fairness, the ‘polluter-pays principle’ can be rather harsh. In many situations, before the era of contaminated sites consciousness-raising (that is, roughly, before 1980 in many developed countries), many individuals or companies contributed to soil contamination without any negative intent. A good example are farmers who paid for waste materials in the 1960s for the purpose of elevating their soils in wetland regions, focusing on land improvement and unaware of any negative site effects due to the presence of contaminants in these waste materials. Today, they may be held responsible for the contaminants in their soils originating from these waste materials. It is a task for the politicians to find practical solutions for these cases. Fairness often demands part or complete governmental financial support.

In many countries, the owner of the estate can be held responsible for the site when no polluter can be identified.

The costs of site investigation, site management and, especially, soil remediation can be substantial. Therefore, many specific arrangements are designed in which the government will at least subsidise the actions that are needed to manage soil contamination, even when the polluter is known. Nowadays, it is generally recognised that an efficient solution of the problem of the many contaminated sites in industrialised countries requires a joint effort between the government and the business community. In other words, the budget must come from the tax payer and private initiatives combined.

1.7.11 Cost-Benefit Analyses

A different way of looking at Risk Management relies on *cost-benefit analyses*, weighing the expected costs against the expected benefits (e.g., Crettaz et al. 2002; Edejer et al. 2003). It is often difficult to quantify the benefits of Risk Management solutions. Grosse et al. 2002, for example, estimated the economic benefits from projected improvements in worker productivity, resulting from the reduction in children's exposure to lead in the United States since 1976. The authors showed that because of falling lead-blood levels, USA preschool-aged children in the late 1990s had IQs that were, on average, 2.2–4.7 points higher than they would have been if they had the blood lead distribution observed among preschool-aged children in the USA, in the late 1970s.

A cost-benefit analysis necessitates that costs and benefits should be expressed in the same units, usually in the unit of money (e.g., euros or US dollars). Generally, a Risk Management solution is beneficial when the value of the benefits is higher than the value of the costs. An optimal Risk Management solution seeks the most optimal (highest) 'value of benefits/value of costs' ratio. Since human health effects are difficult to monetarise, health effects often are expressed as DALYs, that is, a measure for the overall disease burden defined as the sum of the years of life lost due to premature mortality in the population and the years lost due to disability (WHO 2009).

Cost-benefit analyses are also used to evaluate Risk Management projects.

1.7.12 Integration of Human Health and Ecological Risk Assessment

Human Health and Ecological Risk Assessment mainly developed independently of each other. Certainly in the pioneer years of risk-based soil quality assessment, only experts from the respective disciplines were concerned with either Human Health or Ecological Risk Assessment. The role of generalists, who could have promoted integration, was limited during this phase. Another reason for the independent development is that in most countries Human Health Risk Assessment was developed earlier than Ecological Risk Assessment.

In the last few years, there has been a trend with regard to making a case for a stronger link between Human Health and Ecological Risk Assessment (WHO 2001; Suter et al. 2005). The UNEP/ILO/WHO International Programme on Chemical Safety (IPCS) formulated two fundamental reasons for the integration of Human Health and Ecological Risk Assessment, in the framework of the production, use, transport and disposal of chemicals (WHO 2001). First, it improves the quality and efficiency of assessments through the exchange of information between human health and environmental risk assessors. And, second, it provides more coherent input into the decision-making process. Indeed, in several risk-based frameworks different values are used for important input parameters, such as the Koc (organic carbon fraction-based partition coefficient), for Human Health and Ecological Risk Assessment applied to the same site. Moreover, as a third argument, it is wise to focus on the same degree of conservatism/precaution in a Human Health and Ecological Risk Assessment for the same site.

Actually, there is an important fourth reason for the integration of Human Health and Ecological Risk Assessment, and that is to improve the balance in terms of politically defined risk levels. The basic idea of the more integrated Risk Assessment framework proposed by the WHO (2001) is to treat the relationships among Risk Assessment, Risk Management, stakeholder input, and data-collection activities in a general, parallel and concurrent way.

1.7.13 Harmonisation of Risk Assessment Tools

Since the beginning of risk-based contaminated site management in the early 1980s, a large number of *Risk Assessment tools* have been developed in many countries. As a consequence, many Risk Assessment tools exist for the same purpose. By ‘Risk Assessment tool’ is meant a model, regression equation, table, protocol, graph or document, which can be used to determine variables that are used in Risk Assessment. These variables can vary from a ‘supportive’ parameter such as a Kow (octanol water partition coefficient), via the available fraction of an organic contaminant with regard to Ecological Risk Assessment, on up to ‘direct’ input parameters such as measured concentrations in body fluids or tissue.

Although the development of Risk Assessment tools was often based on studying existing Risk Assessment tools, the diversity of tools that are available worldwide for the same purpose is remarkable. This diversity is partly due to different geographical, cultural and social conditions, and sometimes due to differences in political points of view. However, lack of scientific consensus also explains many of the differences. An example of different Risk Assessment tools that serve the same purpose is the procedure to determine concentrations in vegetables for metals (essential for the calculation of exposure due to vegetable consumption). Many countries that have a procedure on contaminated site management derived BCFs (BioConcentration Factors) or regression equations for this purpose. In some Northern European countries, BCF values were adopted from other Northern European countries, but this is an exception rather than the rule. One reason that many countries derived their own Risk Assessment tools might have to do with the fact that the type of vegetables, and the specific genotype of that vegetable, that grow in different countries (and certainly in other climate zones) differs. This is a geographical difference. Another reason is that in different countries, apart from the *possibilities* of growing specific crops, different type of vegetables are grown because of cultural differences or traditions. However, there is no scientific, generally accepted protocol that is used in these countries with regard to the amount of data for each vegetable, quality of the data set, extrapolation margins outside the range of the input data, etc. In Europe, one of the major challenges in Risk

Assessment today is to move towards more consistency in the Risk Assessment tools used by the EU-member states (Swartjes et al. 2009).

Since geographical, cultural and social conditions indeed vary, a worldwide complete harmonisation of Risk Assessment tools is not applicable. However, for the sake of scientific integrity, a stronger convergence of Risk Assessment tools that do not include geographical, cultural, social or policy elements would be favourable (*standardised* Risks assessment tools). Risk Assessment tools that do include geographical, cultural or social elements must be applied with a certain level of flexibility so as to account for these geographical, cultural or social elements (*flexible* Risks assessment tools). Alternatively, guidance could be developed which would describe the requirements for these flexible Risk Assessment tools, which would take into account a higher degree of consistency on the part of the scientific elements of these Risk Assessment tools (Swartjes et al. 2009).

1.7.14 Brownfields

The CABERNET (Concerted Action on Brownfields and Economic Regeneration) network defines Brownfields as ‘sites that have been affected by the former uses of the site and surrounding land; are derelict and underused; may have real or perceived contamination problems; are mainly in developed urban areas; and require intervention to bring them back to beneficial use’ (Oliver et al. 2009). The US Environmental Protection Agency uses a different kind of definition in which the redevelopment or reuse is central, namely, ‘a Brownfield site means real property, the expansion, redevelopment, or reuse of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant’ (US EPA 2009). In Fig. 1.10 an example of a Brownfield in San Francisco, California, USA is shown as an aerial view.

Since the general approach for contaminated site management would imply long development time frames, large uncertainties and an over-proportional budget, a specific Risk Management approach is followed for these Brownfields. It is often claimed that Brownfield redevelopment is primarily financially driven. The general idea, however, is to make site redevelopment profitable, while at the same time protecting human health. Therefore, economic and socio-cultural factors are given greater weight than in ‘normal’ cases of soil contamination.

O’Reilly and Brink (2006) developed a simple screening procedure for Brownfield sites in New York State, USA, in which they classify human health risks in three categories on the basis of the concentration and toxicity of the contaminants, the location of the contaminant, the exposure route (oral, inhalative or dermal) and the type of site user (construction/utility worker, residents, industrial employees, visitors/shoppers). A popular, but very informative overview of Brownfield revitalisation, including examples from the city of Stuttgart in Germany, the Nantes metropolis in France, the cities of Tilburg and Hengelo in the Netherlands, the Medway Council and Torfaen County Borough Council areas in the UK, is given

Fig. 1.10 An aerial view of the San Francisco Bay area, with a Brownfield in the foreground (photo: San Francisco Bay regional water quality control board; reproduced with permission)

in REVIT (2008). A detailed account of the management of Brownfield sites is given in Nathanail (Chapter 20 of this book).

1.7.15 Risk Perception and Risk Communication

Risk Assessment and certainly Risk Management interacts with the daily life of the general public. The life of individuals who do not have knowledge of the effects of contaminants or of the fate and transport processes in soil are affected by contaminated sites, when they live or work on it or are in any other way associated to this site. There have been many cases in which contaminated sites raised enormous concern in the society. See Fig. 1.11, for example, which shows a notice board at which a connection between a landfill and an increased risk for cancer is presumed, in the Silvermines area in Ireland, in 2002. The general public has a much more intuitive approach towards contaminated sites than the experts have. Grasmuck and Scholz (2005) found that humans with higher scores in self-estimated knowledge tended to provide lower risk judgments, were less interested in further information, showed low emotional concern, and thus displayed higher risk acceptance.

The intuitive approach under laymen led to the consequence that the soil compartment does not really have a good reputation. For the general public, soil is a dark place, where some obscure organisms live (if any) and lugubrious decomposition processes take place. You cannot see what is happening in soils (‘and maybe that is just as well’). A much more sophisticated view on soils relates to the soil as

Fig. 1.11 A notice board at which a connection between a landfill and an increased risk for cancer is presumed, in the silvermines area in Ireland, in 2002 (photo: F. Swartjes)

the source for plant production, and hence food production, and drinking water. On the other hand, many people, encouraged as a result of media attention, blame the soil for contaminating their food and drinking water.

Of course, perception of the contaminated sites problem is as varied as human character. Generally speaking, acceptability is less if the associated possible diseases are less known, manifest themselves in the longer term, and when cancer is involved (a small chance for cancer is often perceived as worse than a huge chance for another serious disease). Lima (2004), who investigated the Risk Perception of people living near an incinerator in Portugal, demonstrated that Risk Perception was initially more acute for persons living closer to the incinerator. After a while, however, the persons living closer to the incinerator showed a habituation effect. They developed less extreme attitudes and a lower estimate of the risk.

When laymen are confronted with contaminated sites, they often associate the contaminated site with the diseases that the contaminants *could* generate. Naturally, humans are often afraid of anything dangerous that they cannot comprehend nor control. It is much easier to accept something, even serious adverse effects, that humans can understand and even more so when they are able to control these effects.

For these reasons Risk Communication in an extremely important process. Typically, risk assessors often experience a situation where it is difficult to convince stakeholders of the fact that the risks are acceptable when contaminants are present, even though these are usually at low levels, simply because the diseases that these contaminants *could* generate are known. A well-known example is the presence of asbestos, maybe the best known carcinogenic contaminant in the environment due to intensive media attention, in the soil of a residential garden. The realisation alone that asbestos has been found in the immediate living environment could be cause for panic, independent of the amount, type and condition of the asbestos, let alone the risks involved.

For obvious reasons, the credibility of the risk communicator is of the utmost importance. Stakeholders' participation and intensive communication is crucial, for the purpose of putting the risks into realistic perspective and of supporting stakeholders in making the right decisions. Today, consortium building is therefore an important activity at the start of any contaminated site.

1.8 Approaches Towards Contaminated Site Assessment and Management

1.8.1 Evolution

Since the discovery of soil contamination in the late 1970s, the approach towards soil contamination has undergone a major evolution. Several reasons have contributed to this evolution, for example, the growth in understanding of the risks related to contaminated sites and of procedures for managing the risks. Moreover, the enormous increase in the number of contaminated sites that have been detected made more practical approaches necessary. A very important development in many developed countries is the more integrated approach of contaminated site management with spatial planning. And, finally, the public attitude towards the environment as a whole is constantly evolving. In recent years, the *concept of sustainability* has been advocated in many countries for the state of the environment as well as for activities that impact the environment, including soil quality.

1.8.2 Multifunctionality

In the pioneer years of contaminated site management, contaminated sites were considered to be an incomprehensible threat beyond human control, and for which the adage was rather straightforward: eliminate the whole problem in such a way that every kind of use of the site is possible. This *multifunctional approach* was advocated in many countries in the world and seemed economically feasible as long as the number of sites was limited. The advantage of the multifunctional approach was that no elaborate administrative procedures were needed for keeping an account of the possibilities for and restrictions on the use of a site. Moreover, no complicated Risk Assessment procedures were needed, since risks were, for all intents and purposes, reduced to zero. Regulators did not have to bother about acceptable risk levels.

However, the multifunctional strategy of site management did not free regulators from aftercare activities and costs. 'Dig-and-Dump', that is, removing the contaminated soil from the site and depositing it in landfills, was a popular remediation technology in the early days of contaminated site management. For contaminated groundwater ‘Pump-and-Treat’, that is, extracting contaminated groundwater from the aquifer and cleaning it in water treatment facilities, was often applied. But for bigger sites ‘Dig-and-Dump’ was often too comprehensive and expensive. And extraction of contaminated groundwater did not always explicitly solve the problem, since it resulted in contamination of clean groundwater through desorption of contaminants from the solid phase of the aquifer. For these reasons, mainly financial, technologies that required long-term aftercare were usually applied. An example of such a technology is the *boxing in* of contaminated parts of an aquifer by means of sheet pile walls.

1.8.3 Fitness-for-Use

A much more (cost-)efficient alternative for contaminated site management is the concept of *Fitness-for-Use* (or: Fit-for-Purpose or Suitable-for-Use). This concept implies that the assessment and management of the contaminated site relates to a specific type of land use. This could either be the present or the future land use. The latter is often an option when an alternative type of land use would fit the present soil quality better. Since the late 1980s, the concept of Fitness-for-Use has gained in popularity and has gradually become the leading concept in most countries.

The advantage of Fitness-for-Use is simply that in most cases less strict requirements can be applied. This is much more efficient in terms of the time frames needed for Risk Management activities and costs. Besides that, for many scientists, consultants and regulators, but also for the general public, Fitness-for-Use is a rather logical concept. The idea behind this conception is that, such as most common things in life, things need to be suited for a specific, appropriate purpose. A garage, for example, needs to be suitable for parking a car and not as a playground for children.

The disadvantage of the Fitness-for-Use approach is that aftercare is often needed. Humans can live, work or recreate at a specific site, without experiencing unacceptable human health effects. And when the soil ecosystem is considered as a protection target, the soil ecosystem can be sufficiently protected under specific conditions. However, contaminants might threaten clean groundwater through leaching and migrate to places with a more sensitive land use for human beings or the soil ecosystem. Another drawback of the Fitness-for-Use concept is that intensive administration procedures are needed in order to keep an on-going account of the state of the soil contamination and of the restrictions for the use of a site. Moreover, compared to the multifunctional approach, intensive investigations using Risk Assessment procedures and defining appropriate Risk Management solutions, often including remediation plans, are needed.

In the framework of the concerted action known as *CLARINET* (Contaminated Land Rehabilitation Network for Environmental Technologies) a concept of risk-based soil quality management has been advocated so as to be able to guide the Fitness-for-Use, called *Risk-based Land Management* (*RBLM*).

1.8.4 A More Pragmatic Approach

1.8.4.1 Mentality Change

Since the mid 1990s a significant change in mentality in terms of contaminated site management has taken place in most developed countries. Gradually, the general idea has settled in that the present philosophy on contaminated site management was solid, but too rigid and not efficient enough to make the desired progress. Desired progress, in this perspective, is mainly expressed in terms such as percentage of the total contaminated site load that has been restored. Mainly within the larger municipalities, contaminated sites had resulted in stagnation in building activity.

As a response, the basic concept of management of contaminated sites has evolved into the adage ‘environmentally acceptable and financially feasible’. Fitness-for-Use and cost-efficiency have become important boundary conditions in contaminated site management. It has become widely accepted that not every risk means an *unacceptable* risk. Therefore, the remediation objectives, at least for immobile contaminants in the upper soil layer, need not relate to ‘no risk’, but to ‘an acceptable risk’ for the specific land use (land use-specific remediation). Moreover, it was widely accepted that more cost-efficient methodologies should become available.

At the same time, many governments have developed procedures for financial support, such as sharing financial risks, the provision of subsidy grants, co-financing structures, tax benefits and ‘green’ investments.

1.8.4.2 Natural Attenuation

The concept of using biodegradation as a Risk Management solution, which generated enormous interest in the 1990s and could count on broad support from decision-makers, initiated a very important innovation, that is, *extensive* in situ *remediation technologies* (bio-restoration). Although this is the oldest remediation technology, since organisms have been breaking down organic contaminants ever since contaminants were present in the soil, it was not greatly accepted in the early days of contaminated site management. The reason for this was the relatively long time span that was needed for complete restoration of the site. Moreover, it was difficult to predict the progress of this kind of remediation. However, this change in mentality has come up with the general idea that ‘contaminants that have been in soil for many decades need not necessarily be removed within a time-span of months on up to a few years’.

This insight led to numerous investigations, mainly in the second half of the 1990s, for the purpose of understanding the processes better along with improving efficiency and predicting the time span needed. The process is now called *Natural Attenuation* (aka: intrinsic remediation), which often also includes dilution by transport processes such as molecular diffusion and hydrodynamic dispersion. Natural Attenuation is often considered as a Risk Management solution rather than as a remediation technology. Despite sometimes high starting costs, the overall budget for this Risk Management procedure is relatively low. Besides this, it results in a minimal disturbance of the natural conditions at the site. Today, these extensive remediation technologies, often in combination with ex situ techniques such as removal of the source, are widely accepted. As specific applications, ‘bio-screens’ are used; these are zones with an active, often stimulated, degradation at strategic positions in the soil system, or Funnel-and-Gates techniques, in which contaminants are led to zones with an active degradation. The basic principle is: ‘use the natural self-cleaning capacities of the soil as much as possible, stimulate natural conditions when necessary and use ex situ remediation technologies only when strictly needed’.

1.8.5 Market-Oriented Approach to Site Development

Construction work, certainly within urban areas, is big business. Ironically, most of the interesting locations for residential developments often coincide with contaminated sites. Several former industrial sites, which were created close to the former limits of the municipality, now lie within the expanded cities. These sites, although often contaminated, offer excellent possibilities for residential development. Other examples for residential development are former harbour sites, and former waterfront storage depots or warehouses, which offer exclusive housing and recreation opportunities. Again, these sites are often contaminated. An interesting adage from site contractors involved in building activities such as developing residential areas is: ‘turn a threat into an asset’. Risk Management of contaminated sites in these areas could result in a profitable rise in value of the site for different building purposes. Moreover, consultancies have proven to be experts in finding creative solutions for risk reduction.

More and more it has become accepted fact that contaminated site management is part of the integral complex package of site development (see Fig. 1.12 as an example, which shows a former with DDT (DichloorDifenylTrichloorethaan) and HCH (HexaChloroCyclohexane) contaminated site in Wuxi, China, which has been developed into a residential area after remediation). Traditionally, contractors had to adapt the physical state of a site through grading, providing drainage and guaranteeing the supporting foundation of buildings by the use of piles. One could argue that maybe contaminated site management is just another aspect of making the site suitable for building activities. Contaminated site management is just another aspect that should be included in a cost-benefit analysis. As a consequence, market-driven financing is contributing more and more to solving the problem of contaminated sites.

The development of a more market-oriented approach has gone hand in hand with the mentality change towards a more flexible way of contaminated site management and a more intensified use of Natural Attenuation techniques (Section 1.8.4.2). As a consequence, ‘the market’ has taken on the responsibility for a cost-efficient risk reduction at many contaminated sites.

Schelwald-Van der Kley et al. (Chapter 24 of this book) describe the philosophy on cost-efficient Risk Management solutions of industrially contaminated sites, at the same time discussing the protection of human health, ecology and groundwater.

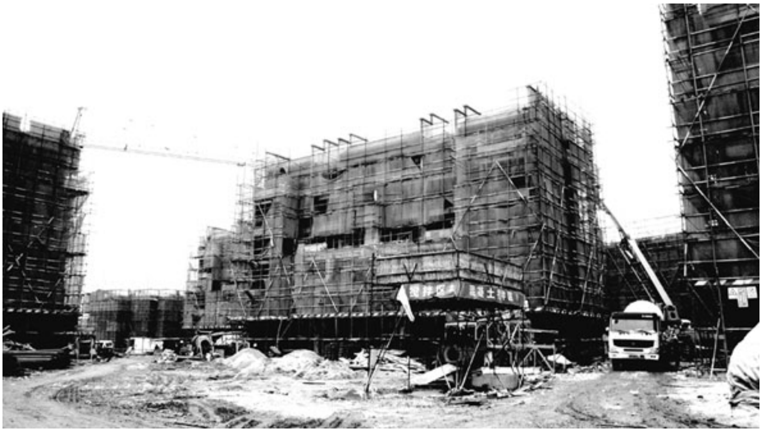


Fig. 1.12 A former with DDT (DichloorDifenylTrichloorethaan) and HCH (HexaChloro Cyclohexane) contaminated site in Wuxi, China, which has been developed into a residential area after remediation (photo F. Swartjes)

1.8.6 Integrated Approaches

Some individuals involved in risk-based soil quality assessment believe that the main purpose of the term 'integrated', mainly in the 1990s, was to 'dress up' political reports and letters. Although this opinion does not reflect reality, it is a fact that plans for and ideas about integration always overruled concrete application. Today, several variations on integration offer immense benefits in contaminated site management.

1.8.6.1 Interdepartmental

In all countries in the world, various ministries have a relationship with soil quality assessment and management. Political themes that have a relationship with soil include environment, agriculture, water resources, nature protection, and spatial planning. A balanced *interdepartmental approach*, however, would practically be impossible. Therefore, it is essential that laws and acts that influence soil policy do not permit actions to conflict, or, still better, that they actually strengthen each other. The same conclusions hold for national versus international regulations. In Europe, for example, many environmental acts, such as the Water Framework Directive, overlap with national regulations.

1.8.6.2 Spatial Planning

Traditionally, spatial planning is a process with a two-dimensional scope, that is, it is related to the arrangement of the soil surface, usually on the scale of a region. Since the late 1990s, the idea of including the third dimension, that is, soil aspects, into spatial planning, has become a point of interest. The reason for this is that the soil, that is, both the upper layer and the groundwater, impose limitations as well as opportunities for functions at the soil surface. An example of these limitations concerns the establishment of a new estate district or an industrial park on soils with limited bearing capacity (e.g., peat soils). An example of an opportunity is the planning of a nature reserve, in combination with additional water storage possibilities in wet agricultural areas, in case of high water recharges, in anticipation of climate change. Although humans are capable of changing the environment, including many soil factors, in accordance with their own requirements, these examples show that this is not always possible, or at least not without substantial additional costs. Therefore, several soil issues, including soil quality assessment and management must become part of any integrated spatial planning process.

Moreover, in densely populated regions in the world, sub-surface construction works have gained in popularity. From this perspective it becomes unavoidable to include the third (depth) dimension, i.e., the soil quality, in spatial planning.

1.8.6.3 Chemical, Physical and Biological Soil Quality Assessment

Traditionally, since contaminated sites have become a political issue, soil quality is approached from a chemical perspective, that is, focused on contaminants in soil. During the last few years the philosophy of considering overall soil quality, that is, chemical, physical and biological soil quality, gained in interest. A concrete example is the determination of so-called Soil Ambitions for local soil quality in the Netherlands (Otte et al. 2009). These Soil Ambitions can be assessed at the level of municipalities, with the use of a so-called Route planner. Although the chemical component of the assessment is the most mature, physical elements (such as sealing, or bearing power), or biological elements (such as Biodiversity) could also be included in Soil Ambitions.

1.8.6.4 Environmental, Socio-Cultural and Economic Assessment

Another interesting integration is to combine risk-based soil quality assessment with social and economic factors. In a way, economic factors always have been included, since no Risk Management activities have had access to unlimited financial resources. Today, however, cost-efficiency has become an important element of modern risk-based soil quality assessment. Socio/cultural factors have also implicitly played a role of some importance. This is reflected, for example, in the different approach to contaminated site management in urban areas than that applied in rural areas.

In the procedure for dealing with Brownfields, environmental, socio-cultural and economic factors are assessed and weighed in a systematic way.

1.8.6.5 Life Cycle Assessment

Life Cycle Assessment (LCA; aka: *Life Cycle Analyses* or *Life Cycle Impact assessment*) is the holistic evaluation of the environmental impacts of a specific product or service, from the time that the product or service comes into existence until the product is deposed of or the service is ended. For products, this usually relates to the overall environmental impacts of the product in all stages, that is, the raw material production, manufacturing, distribution, use in practice and disposal. In all these stages there is a possibility that soils are impacted. This impact relates to contaminants from the product under investigation, but also from other products that are needed for manufacturing (e.g., degreasing liquids), and that are released during distribution (e.g., from fuels during road transport) or during use (e.g., from detergents during use). In a Life Cycle Inventory the pathways of the contaminants, including the immissions to the soil compartment, are investigated. Finally, the impact is investigated in the Life Cycle Assessment, which overlaps with risk-based soil quality assessment. Risk Assessment could be an important process, and is in fact one of the steps of the more integral approach of Life Cycle Assessment.

Life Cycle Assessment can also support the selection of Risk Management solutions. Suèr et al. (2004), for example, described nine case studies in which Life Cycle Assessment was used to evaluate alternative remediation technologies. As criteria they use the spatial scale that is affected by the remediation, the time scale at which positive and negative effects reveal themselves and secondary processes such as the production of tubing, electricity use, materials used for treatment of contaminated groundwater, the production of iron fillings for reactive walls, and of active carbon, nitrate, and hydrogen peroxide for bioremediation purposes.

1.8.7 Technical Approaches

1.8.7.1 Risk Assessment Methodologies

During the last few years, many easily accessible Risk Assessment methodologies have become available. These methodologies often concern spreadsheet-like models in a Windows environment. Many of these models are readily available, for example, on the Internet. The huge advantage of this is that the Risk Assessment process has gained in popularity and is followed more often, which generally improves the procedure on contaminated site management. However, there also is a serious disadvantage, since engineers can use these models without much knowledge of Risk Assessment. It is important to realize that these models may not be used as black box models. Even more disturbing is the fact that these procedures, though in fact any Risk Assessment methodology, are easy to manipulate according to the wishes of a specific stakeholder. Therefore, any Risk Assessment needs to be accompanied by some kind of certification process. Usually regulators are primarily responsible for the objectivity and quality control of the Risk Assessment.

Another matter of concern is that individuals who are not expert often believe in models that they do not understand. Ironically, it sometimes happens that laymen trust models more when they appear to be more complex. It is the responsibility of the risk assessor and the regulators to provide transparency in the possibilities and the limitations of Risk Assessment methodologies.

An interesting and difficult issue in Risk Assessment methodologies concerns the balance between uniformity and flexibility. Uniformity is an important regulatory aspect. Two different risk assessors must come to the same conclusion about the risk appraisal of a specific site, if only from the perspective of fairness. In general, uniformity improves when Risk Assessment procedures are more rigid, with more fixed input parameters. However, a higher degree of freedom in the use of Risk Assessment methodologies stimulates the incorporation of more site-specific information, which improves the assessment, at the expense of uniformity. It is the task of both the scientists and the regulators to strive towards an optimal balance between uniformity and flexibility when making Risk Assessment methodologies accessible.

Another challenge is to find a good balance between the scientific foundation and the possibilities for pragmatic applicability. A scientist might claim that many difficult theories cannot be put into a practical format. However, practical methodologies which are scientifically not completely mature are often better than no methodologies at all. When this kind of methodology is used, the scientific limitations need to be made transparent. These scientific limitations are the basis of the interpretation of Risk Assessment results, at the same time offering the possibilities for improvement of the methodology. It takes courage, however, to follow this practical approach, since scientists thereby make themselves vulnerable.

1.8.7.2 Conceptual Model

Every Risk Assessment is somehow related to a source-pathway-receptor approach. With regard to risk-based soil quality assessment, it often pays off to start with a Conceptual Model, especially in the case of contaminated aquifers. Such a Conceptual Model gives a (usually visual) presentation of relationships between the source, all pathways involved and the receptor. A cross-section of the contaminated site is the most common format of a Conceptual Model. See Fig. 1.13 in which a Conceptual Model for groundwater-to-indoor-air mass flux analysis is shown, as an example.

In fact the Conceptual Model represents a two-dimensional contamination pattern and includes all relevant pathways involved. It relates to the migration to other compartments, for example, from soil to aquifer or to the migration within a compartment, for example, the migration of a contaminant plume within the aquifer. It also represents migration of contaminants into contact media, for example, upward transfer of volatile contaminants from the upper aquifer into a building. The source generally relates to a location, or locations, in soil that are contaminated. The receptor is a specification of the protection targets. It can relate, for example, to humans living on the site, or to the soil ecosystem in a downstream nature reserve that might be threatened by a lateral and upward flow pattern in the aquifer and upper soil layers.

A Conceptual Model might serve two important purposes. First, it supports a systematic investigation of all possible pathways, and subsequently helps to identify all necessary Risk Assessment tools. Second, it makes the whole Risk Assessment

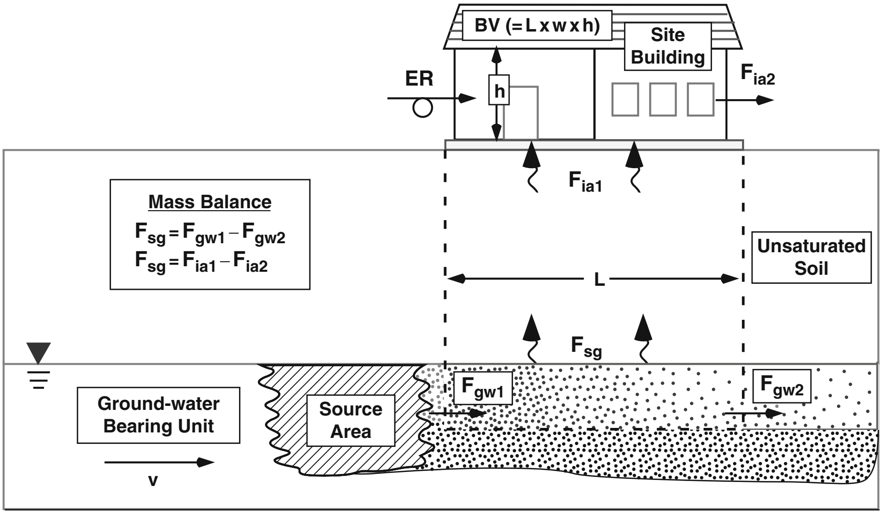


Fig. 1.13 A Conceptual Model for groundwater-to-indoor-air mass flux analysis, as an example of a Conceptual Model (source: McHugh et al. 2003; reproduced with permission). Fgw1 = mass flux through groundwater at upgradient edge of building, Fgw2 = mass flux through groundwater at downgradient edge of building, Fsg = mass flux from groundwater to soil gas under building, Fia1 = mass flux from soil gas to building, Fia2 = mass flux from building to outdoor air

process accessible, even to non-experts. For this reason, a Conceptual Model is an excellent basis for discussions in an early stage of any Risk Assessment project, along with the involvement of all stakeholders.

1.8.7.3 Tiered Approach

The most elegant way of dealing with Risk Assessment is in the form of a tiered approach. In such an approach several assessment steps (*tiers*) are described. In each tier, an assessment is performed with generally two possible outcomes: either a judgment of the absence of unacceptable risks can be given, and the total assessment is finished, or unacceptable risks cannot be excluded and the assessment has to be followed into the next tier. When the presence of unacceptable risks cannot be reputed in the final tier, unacceptable risks cannot be excluded, the total assessment is also finished and Risk Management needs to follow the tier-based Risk Assessment. Given the nature of a tiered approach, in each step the assessment becomes less conservative, is based on more site-specific information and, hence, is more complex, time-consuming and often more expensive. The philosophy behind this is: *simple when possible (only the first tier) and more complex when necessary (higher tiers)*. A tiered approach is an efficient way of risk assessing, without compromising scientific integrity.

A tiered approach often starts with a generic Risk Assessment, that is, comparing measured value with Soil Quality Standards (screening levels). The purpose could be, for example, to separate risk factors into two classes, namely, ‘no unacceptable risks’ and ‘further research needed’. In this further research stage, fine-tuning of risks is performed in one or more higher tiers. Another option is the tripartition of risk classes in the first tier, namely, ‘clearly no unacceptable risk’, ‘clearly an unacceptable risk’ and an intermediate class. Obviously, the fine-tuning in higher tiers needs to occur in the last (intermediate) class.

1.8.7.4 Weight of Evidence

There is no uniform definition of *Weight of Evidence* (WOE), but it refers to multiple lines of investigation, for which the results are combined to improve the reliability of an assessment. The principle behind the Weight of Evidence approach is that when several uncertain results are combined, the overall result is less uncertain. Weight of Evidence is a popular tool used in forensics (e.g., Balding 2005). Since Risk Assessment typically is associated with uncertainties, WOE is often used, although the term ‘evidence’ usually must not be taken literally. Most approaches concern a more qualitative way of combining the results of the multiple lines of evidence. Smith et al. (2002), for example, described a quantitative Weight of Evidence approach to predict the risk at potentially contaminated sites on the Great Lakes in the USA.

Weed (2005) described three kinds of ‘evidence’ in the Weight of Evidence approach. The first is metaphorical where Weight of Evidence refers to a collection of studies. The second is methodological where Weight of Evidence points to different methodologies. In the third, the author referred to theoretical evidence where Weight of Evidence serves as a label for a conceptual framework. As a practical recommendation, it was stated that the WOE concept and its associated methods should be fully described when used.

1.8.7.5 Decision Support Systems

In standard works (e.g., Finlay 1994) a *Decision Support System* (DSS) is often rather broadly defined as a computer-based system that aids the process of decision-making. In the field of risk-based soil quality assessment, a Decision Support System is a methodology, often in the format of a user-friendly computer program, for which the generated outcomes are linked to concrete (regulatory) consequences for contaminated site management. It generally combines Risk Assessment tools with political positions. Decision Support Systems often have a relatively large degree of objectivity and are, hence, rather rigid to use.

DSSs have some advantages. First, using DSSs is more efficient than an ‘open’ site-specific Risk Assessment, since many of the choices on Risk Assessment tools (models, equations, input parameters) have already been made. For the same reason, it improves uniformity. Second, it facilitates communication, at least when DSSs have been described in detail.

An example of a DSS is the procedure used in the Netherlands (Dutch Soil Protection Act) to determine the urgency of remediation, based on site-specific risks for human health, the ecosystem and contaminant migration in the groundwater (Swartjes 1999).

For more detailed frameworks for contaminated site management, see Vegter and Kasamas (Chapter 23 of this book) from the viewpoint of policy makers and regulators; Schelwald-Van der Kley et al. (Chapter 24 of this book), from the viewpoint of the industry as an important stakeholder; and Nathanail (Chapter 25 of this book) with regard to the management of Brownfields.

1.9 Sustainability

Sustainability relates to a specific state that is able to continue indefinitely. A very workable and often-cited definition of a sustainable development is the definition of the General Assembly of the United Nations (United Nations 1987), often referred to as the Brundtland Commission, that is, ‘a development that meets the needs of the present without compromising the ability of future generations to meet their own needs’. The good thing about this definition is that it focuses on future continuation (future generations must meet their own needs), but also offers a realistic perspective for the present (meeting the needs of the present).

This view of sustainability accounts for the major pitfall for contaminated site management, that is, the sole focus on short-term benefits, since spreading the benefits over longer time spans is a luxury that is not always permitted. In the so-called ‘Rio declaration on environment and development’, the outcome of the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro (often called the Earth Summit) (United Nations 1992), the principle of sustainability was further stressed on the widest international scale possible.

The *Cradle to Cradle Concept*, which was introduced by a designer-chemist duo (Braungart and McDonough 2002), goes one step further. This concept is not limited to preservation, but rather focuses on improvement in the quality of life. The motto ‘the waste of today must be the food for tomorrow’ illustrates the idea of the Cradle to Cradle Concept. Although the authors aroused a lot of scepticism—according to criticasters the idea was too general and the authors did not specify the road map towards this ideal world perspective—their philosophy opened the way to many discussions on improving long-term care for the planet, including the world’s soils.

With regard to contaminated site management, it seems wise not only to focus on the time factor, but also to incorporate the undesired shifting of problems to other places. In fact, the character of contaminated sites (more specifically, contaminated groundwater), that is, a dynamic system with migration of contaminants, forces the risk assessor to account for off-site risks. As a consequence, the wider definition of sustainability used in this book is: a development that meets the needs of the present *at a particular site* and without compromising (or with improvement on) the ability of future generations to meet their own needs, and *also in the wider surroundings*. From this perspective the plea for a regional-scale approach (see Section 1.7.8) is justified.

Parallel with the increase in anthropogenic pressure on the environment, including the soil, sustainability has gained enormously in popularity. The basic reason for this is that history has proven that successful cultures have eventually been wiped out by their own success. This phenomenon seems to relate currently to most developed countries, since human presence, activities and production have already impacted many natural resources which were in equilibrium for long time frames. A highly contemporary example is the increase in the immission of substances into the atmosphere, leading to climate change. It seems that without sensible human interference the indefinite maintenance of a stable climate suited for human survival is not guaranteed.

Like the climate, soils are also strongly impacted by human activities. Obviously, human interaction with soil has a huge economic potential and should not be avoided. Sustainable soil management, however, implies that human interventions should adapt to the rhythm of natural soil processes. *A healthy soil ecosystem*, with appropriate resilience and recovery abilities, is generally considered to be an indicator of sustainable soil quality. Many processes that lead to soil contamination typically are not sustainable, since the soil ecosystem is often not able to respond properly. Without regulations, important soil processes will not continue optimally. More than that, several human activities might lead to an irreversible elimination of specific soil organisms and, hence, block ecological processes, more or less the opposite of sustainability.

An important aspect of sustainable contaminated site management relates to *prevention* of soil contamination. Jenck et al. (2004), for example, claimed that 'industrial sustainable chemistry', with process design and new equipment leading to minimal immissions, is seen worldwide today. They illustrated this statement with several current industrial case studies.

1.10 Actors Involved

1.10.1 Decision-Makers and Regulators

Ideally, contaminated site management needs to be regulated in acts or laws. Therefore, decision-makers must form the basis of any (national or regional) framework for contaminated site management, which is generally based on risk-based soil quality assessment nowadays. Many European countries, the USA, Australia, Canada, and some Asian, South American and African countries have enacted legislation on contaminated sites. The structure of involvement and responsibilities of the government differ among countries, not least of all because the size of the countries differs. Bigger countries, such as Canada or Spain, for example, have legislation at the provincial level. There also are countries that have national legislation with specifications at the provincial level.

A typical task for decision-makers, in close cooperation with scientists, is the selection of protection targets and the determination of protection levels. Another important task for decision-makers concerns the indication of the level of desired conservatism/precaution, for example, protection of the average human being, or protection of all human beings. Usually, the protection levels and the level of desired conservatism need to be determined in agreement with other existing laws. Although decision-makers play a crucial role in the definition of boundary conditions and degree of desired conservatism, it is essential that they unbiasedly commit themselves to the scientific part of the Risk Assessment and Risk Management.

Except for following fixed protection levels, decision-makers have other options for soil quality appraisal. These options are mainly implemented when defining the end goal of Risk Management, for example, through remediation. First, they can relate the acceptable soil quality to background concentrations, independent of associated risks. From the point of view of efficiency, an elegant procedure is to base the end goal on the *ALARA* (*As Low As Reasonably Achievable) principle*, using *Best Available Technologies (BAT)*. This means that risk reduction is performed up to a concentration level, possibly with a maximum value that is ‘good enough’, for which the costs ‘are reasonable’. Obviously, the determination of ‘reasonable costs’ in relation to the improvement of soil quality is a subjective process. A special level of protection, not directly related to risks, is based on the *stand-still principle*, generally applied to assess the appraisal of groundwater quality or soil material transposition. According to this principle the soil quality may not deteriorate, in other words, contaminant inputs must equal contaminant outputs.

An approach that decision-makers can choose to follow is the *precautionary principle* (Raffensberger and Tickner 1999), especially in case of lack of scientific consensus. This principle is morally and politically based, and states that if an action or process has adversely impacted human health, the burden of solving the problem falls on those who caused the problem. At the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro (often called the Earth Summit) (United Nations 1992), the precautionary principle was advocated with regard to environmental protection. When applying the precautionary principle to contaminated sites, decision-makers can decide, for example, in the case of doubt about the risks involved, that the polluter is responsible for (financing the) elimination of the contaminants.

Along with the definition of protection levels, the decision-makers also have the duty to communicate them. It is often asserted that the political underpinning of these protection levels is based on coincidental aspects and very often not made transparent. In the most extreme case, protection levels that the scientists (arbitrarily) derive are implemented in soil quality assessment frameworks, since ‘they are the only levels available’ in the view of the decision-maker. Moreover, there is a risk that a specific degree of conservatism ‘slips’ into the Risk Assessment, since the scientists use the parameters that are available, without initially focussing on a specified degree of conservatism. It is important that all policy decisions are clearly specified in reports underlying, for example, Soil Quality Standards, Decision Support Systems, etc.

1.10.2 Scientists

Successful management of contaminated sites, whether on a local, regional or national scale, relies on understanding and applying a large and multi-disciplinary knowledge base that straddles the natural, physical, engineering and social sciences within a practical, commercial and political context (Pollard et al. 2002b). Scientists play a principal role in Risk Assessment procedures. Human Health Risk Assessment, Ecological Risk Assessment, Groundwater-related Risk Assessment and Food Safety-related Risk Assessment are based on a number of scientific Risk Assessment tools; these are an equation, a description, a database, a model, an instrument, a protocol, or a table. Basically, the scientists are responsible for the objective development and application of Risk Assessment and Risk Management tools.

It is difficult to profile the ideal risk assessor, since in an overall Risk Assessment several technical disciplines are needed. Moreover, specific social competencies are required. For this reason, most Risk Assessment projects are performed by a multi-disciplinary team, rather than by one single person. The basic kinds of expertise that are essential for risk-based soil quality assessment are soil science, (soil) chemistry/geochemistry, (geo)hydrology, toxicology and biology. Other kinds of expertise that support the Risk Assessment and Risk Management process are mathematics, information technology, statistics, geology, geography, and law. Moreover, any Risk Assessment team benefits from people with a social science background; these are people with communication knowledge or skills. Moreover, people with creative qualifications can make significant contributions to original site-specific Risk Management solutions, with a good balance between quality and cost-efficiency. In rare cases, technical expertise and communicative and creative qualities are combined in the same person.

1.10.3 Decision-Makers Versus Scientists

Procedures for risk-based soil quality assessment are based on scientific Risk Assessment tools and policy decisions. Since these elements are interwoven, the derivation of these procedures concerns a co-production between scientists and policy makers. As early as 1983, a good partnership between science and decision-makers was seen as an essential element of Risk Assessment (US National Research Council 1983). In most countries, however, there is no clear boundary between the tasks of policy makers and scientists. However, the relationship between decision-makers and scientists is crucial for a successful and efficient procedure for contaminated site management. Scientists need the policy makers for the definition of the problem formulation and the definition of the boundary conditions for Risk Assessments. Decision-makers, for their part, need the scientists to explain the possibilities for and exact meaning of protection levels, and the uncertainties involved, and to elaborate on technical procedures, models, protocols and related Soil Quality Standards. Although decision-makers and scientists usually have a common background, that is, academic study in a related technical field, the interest and mentality of both groups is often very different. Moreover, optimal co-ordination requires frequent communication, which is hampered by a physical separation of working places. Even more importantly, mutual communication is not implicitly productive and the significance of it, therefore, is often underestimated.

Souren (2006) investigated the role of scientists and policy makers in the derivation of Soil Quality Standards in the Netherlands in the period 1971–2000. She concluded that in general policy makers do have sufficient scientific knowledge to understand the scientific procedures used to derive Soil Quality Standards. At the same time, the author concluded that most scientists understand policy-making.

1.10.4 The Risk Assessor

Risk Assessment is often said to be the scientific part of the contaminated site management framework and an objective procedure. That might be true for a large part. However, compared to other scientific disciplines, Risk Assessment involves quite as many subjective decisions. An important example of a subjective judgement is the interpretation of measured soil concentrations, certainly in the case of heterogeneously contaminated sites. One risk assessor might focus on the highest measured value, while another assessor will use the average of all samples in the Risk Assessment. In the latter case, it might make a difference whether this individual decides on the arithmetic or geometric mean, which might differ with a factor of up to 10. Ideally, these decisions are protocolised. However, using margins in many decisions (such as the decision on which value to take for the 'representative soil concentration' in the example shown above) could actually improve the Risk Assessment. For example, in an early tier of Risk Assessment and when the risk assessor's estimate is that risks are acceptable, the risk assessor might deliberately focus on the highest concentration. In that case, the most likely outcome is 'risk acceptable and case closed', on the basis of a limited effort and budget. In other cases, for example, when the layout of the site shows different uses such as green shoulders and paved areas at a business park, the risk assessor might derive an average representative concentration, only valid for the green shoulders. In this case, since no vegetables will be cultivated on the shoulder, this average soil concentration will be good enough to investigate the possibilities for the development of typical shoulder vegetation. Using the margins in an intelligent and responsible way, in fact, makes Risk Assessment a creative and challenging process.

The risk assessor needs to accompany the technical Risk Assessment with three very important activities. First, the risk assessor should always relate the purpose of the Risk Assessment to the choices he or she has to make. Second, the risk assessor should incorporate uncertainties in every step of the Risk Assessment. Third, and last but not least, the risk assessors should make every choice, as well as the exact meaning of the Risk Assessment results, transparent. Moreover, during the whole process, communication with decision-makers and stakeholders is essential.

The quality of a Risk Assessment and the uniformity of the results are of utmost importance for a justified risk appraisal and a sensible way of spending the available resources. Clearly, all stakeholders benefit from sound decision-making, based on systematic and 'smart' Risk Assessment. This 'smart' Risk Assessment requires a knowledge base and creativity. Creativity is hard to steer, but the knowledge level of risk assessors could be approved by, for example, auditing systems, courses and exchange of state-of-the-art Risk Assessment tools. Some countries have auditing systems and most countries that have a policy on soil quality organise courses on Risk Assessment.

1.10.5 Project Managers

Project managers are, of course, of crucial importance in any project. The role of contaminated site project managers, certainly for bigger projects, requires special attention. This role is complex, though interesting, for several reasons. First, a project manager needs to be able to at least have an overview of the scientific knowledge of Risk Assessment tools. Because of the multi-disciplinary character of Risk Assessment, this overview often requires the ability to make a broad scientific interpretation on the part of the project manager. Second, the relationship between scientists and decision-makers or regulators traverses the whole Risk Assessment and even more the Risk Management process. Some projects lack efficiency, and often quality, because of the fact that decision-makers are unable or unwilling to use the scientific knowledge they need for sound decision-making. Analogously, scientists often are not capable, nor trained, to include 'soft' decision-making factors into their investigations, such as decisions on boundary conditions related to the degree of conservatism or specific political requirements in their Risk Assessments. It is the role of a project manager to bring together the political starting points and boundary conditions and the scientific elaboration of Risk Assessment tools, throughout the whole development of the project. Third, the project manager needs to have above-average communication abilities in order to be able to deal with the multi-stakeholder character of contaminated site projects. This has two different purposes, that is, the project must benefit from the knowledge and interest of all available stakeholders, and the stakeholders with less knowledge, such as land owners or inhabitants, need to be informed about developments that often affect their living environment and their well-being. Management involves different roles, both technical and social, which include aspects of work relationships as well as personal relationships. This includes safe-guarding (e.g., of the time schedule, the budget), motivating participants, team building and the optimal distribution of working tasks. It is often claimed that the creation of a good social atmosphere also stimulates the more technical performances of a project.

One interesting issue concerns the role of women in Risk Assessment and Risk Management, and, particularly, in project management. In the 'Rio Declaration on Environment and Development', the outcome of the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro (United

Nations 1992), the vital role of women in environmental management and development was stressed. Many studies addressed and many books have been written about the role of women in project management. The success of a female manager very much depends on factors other than gender, such as personality, age, experience, team structure, complexity and interest in the project. Of course, female managers also have disadvantages compared to male managers. Some men, for example, find it difficult to work under the supervision of a female (Fairhurst 1993), while some women are uncomfortable supervising men (Williams and Locke 1999). But, otherwise, women tend to be more relationship-oriented (Fairhurst 1993). Although it would be beyond the scope of this book to make a full analysis of all these factors, it can be stated that some typical female characteristics are beneficial for team building and successful communication and, hence, for Project Management.

1.10.6 Major International Institutions

Some major international institutions provide crucial information on Risk Assessment and Risk Management related to contaminated site management. Their reports are important for at least two reasons. First, the scientific basis of these reports is generally high. Second, using procedures from these institutions improves general use and harmonisation of Risk Assessment tools, while at the same time large-scale use improves the status of these procedures. Without the pretence of being complete, some crucial institutions are mentioned in this section. Since it is hardly possible within the scope of this book to enlarge upon all the relevant topics these institutions have dealt with and are dealing with and to describe the many crucial reports available, it is simply advisable to browse the Internet sites for information on specific topics. To facilitate this process, the institutions are briefly introduced in the footnotes in this section.

Some of the important international institutions are the World Health Organisation (WHO),[2] the International Programme on Chemical Safety (IPCS),[3] and the European Centre for Ecotoxicology and Toxicology of Chemicals

---

[2] The *World Health Organisation (WHO)* is the directing and coordinating authority for health within the United Nations system. It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries and monitoring and assessing health trends (http://www.who.int).

[3] The *International Programme on Chemical Safety (IPCS)* was developed and structured on the basis of recommendations of the United Nations Conference on the Human Environment (1972). It is a cooperative venture between WHO, International Labour Organisation (ILO) and the United Nations Environment Programme (UNEP) ILO and UNEP. The two main roles of the IPCS are to establish the scientific basis for the safe use of chemicals and to strengthen national capabilities and capacities for chemical safety (http://www.who.int/ipcs/en/).

(ECETOC).[4] Some authoritative national institutions in the USA that relate to contaminated site management are the US Environmental Protection Agency (US EPA),[5] the Agency for Toxic Substances and Disease Registry (ATSDR)[6] and The National Environmental Health Association (NEHA).[7] In Europe, The Joint Research Centre (JRC)[8] acts at the level of the European Union. Some important authoritative national institutions are the National Institute of Public Health and the Environment (RIVM)[9] in the Netherlands, The Environment Agency (EA) in the

---

[4] *ECETOC* is a scientific, non-profit, non-commercial trade association with a mission to act as an independent, credible, peer-reviewed technical resource to all concerned with the identification of research needs and provision of scientific rationale for the assessment of health effects and environmental impact, and thereby to justify the industry's license and freedom to operate (www.ECETOC.org).

[5] The *U.S. Environmental Protection Agency (US EPA)* leads the Unitd States' environmental science, research, education and assessment efforts. The mission of the Environmental Protection Agency is to protect human health and the environment. Since 1970, EPA has been working for a cleaner, healthier environment for the American people. The EPA headquarters are in Washington, DC, but there are many other locations, such as regional offices, regional visitor guides, laboratories and research centres (http://www.EPA.gov).

[6] The *Agency for Toxic Substances and Disease Registry (ATSDR)* is an agency of the U.S. Department of Health and Human Services. The mission is to serve the public by using the best science, taking responsive public health actions, and providing trusted health information to prevent harmful exposures and disease related to toxic substances. ATSDR is directed by congressional mandate to perform specific functions concerning the effect on public health of hazardous substances in the environment. (http://www.atsdr.cdc.gov/about/index.html).

[7] The *National Environmental Health Association (NEHA)* had its origins in the state of California, USA, where it was incorporated in 1937. The original impetus behind the creation of a national professional society for environmental health practitioners was the desire by professionals of that day to establish a standard of excellence for this developing profession. NEHA's mission is 'to advance the environmental health and protection professional for the purpose of providing a healthful environment for all' is as relevant today as it was when the organisation was founded (http://www.neha.org/about/neha.html).

[8] The *Joint Research Centre (JRC)* is a research based policy support organisation, and an integral part of the European Commission. The JRC is providing the scientific advice and technical know-how to support a wide range of EU policies. Their status as a Commission service, which guarantees their independence from private or national interests, is crucial for pursuing our mission. This mission is 'to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies'. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

[9] The *National Institute for Public Health and the Environment (RIVM)* is a recognised leading centre of expertise in the fields of health, nutrition and environmental protection. RIVM's mission is to benefit people, society and the environment, matching thier expertise, knowledge and research with that of colleagues from around the world. The institute works primarily for the Dutch government, but shares their knowledge with governments and supranational bodies around the world. The results of RIVM's research, monitoring, modelling and Risk Assessment are used to underpin policy on public health, food, safety and the environment (http://www.rivm.nl).

UK and Wales[10] and the Flemish Institute for Technological Research (VITO)[11] in Flanders, Belgium.

1.11 Scope of the Book

To sum up the previous sections, this book focuses on the possibilities for investigating whether ‘there is a problem’ with potentially contaminated sites (Risk Assessment) and, if so, how to deal with this situation (Risk Management). The scope of the book is limited to the extent of that part of the earth’s crust that impacts human health and the ecosystem, that is, the water-saturated upper soil layer and the groundwater that is within human reach. The book is primarily focused on procedures for dealing with *terrestrial* contaminated sites, not on surface water or sediments.

The book deals with contaminated sites, either diffusely contaminated or locally contaminated (although since most chapters deal with tools that can be used for any type of contaminated site, this distinction is not always relevant). Physical quality of the soil is not within the scope of the book. Also the risks of radioactive contaminants, endocrine disruptors, microbial contaminants and nanoparticles in soil or groundwater are not included in this book, since they are of a different nature and require a different kind of Risk Assessment. This book does not focus primarily on agricultural practices, therefore, no further attention is given to (the consequences of) (macro) nutrients (nitrogen, phosphorus, potassium, and sulphur). However, Risk Assessment tools are described that can be used to assess the risk for Food Safety.

Since in most countries separate legislation exists for the assessment and control of occupational risks, this book does not implicitly focus on the risks of humans that are exposed to soil contaminants during working activities such as digging or other ground construction works. Many of the tools for Human Health Risk Assessment that are described in this book, however, could in principle be used for this purpose.

---

[10] The *Environment Agency (EA)* is an executive non-departmental public body responsible to the Secretary of State for Environment, Food and Rural Affairs and an Assembly-Sponsored Public Body responsible to the National Assembly for Wales. Their principal aims are to protect and improve the environment, and to promote sustainable development. The EA is going through changing times, both in terms of the physical and business environments. Their new strategy builds on the improvements they have delivered in recent years, but it provides a new framework to guide our work to protect and improve the environment of England and Wales in the challenging climate of the next 5 years (www.environment-agency.gov.uk).

[11] The *Flemish Institute for Technological Research (VITO)* has, in its 18 years of existence, achieved the status of being a key player in the European world of research and development. In the research domains of environment, energy, materials and remote sensing, VITO’s strength has been its pursuit of applied and practical research and development which is relevant for industry and public authorities. VITO expresses its ’vision on technology’ through the recommendations given to clients vis-à-vis technological developments, as well as in the way in which the VITO experts develop new technology and support companies with their innovation (http://www.vito.be).

References

Altfelder S, Beyer Ch, Duijnisveld WHM, Schneider J, Streck Th (2002) Distribution of Cd in the vicinity of a metal smelter: Interpolation of soil Cd concentrations with regard to regulative limits. J Plant Nutr Soil Sci 165(6):697–705

Alvarenga P, Palma P, Gonçalves AP, Fernández RM, de Varennes A, Vallini G, Duarte E, Cunha-Queda AC (2008) Evaluation of tests to assess the quality of mine-contaminated soils. Environ Geochem Health 30(2):95–99

Amrate S, Akretche DE, Innocent C, Seta P (2005) Removal of Pb from a calcareous soil during EDTA-enhanced electrokinetic extraction. Science Total Environ 349(1–3):56–66

Arienzo M, Adamo P, Cozzolino V (2004) The potential of Lolium perenne for revegetation of contaminated soil from a metallurgical site. Sci Total Environ 319(1–3):13–25

ASTM (2009) ASTM E2081 – 00(2004)e1, Standard guide for risk-based corrective action, ASTM E2081 – 00(2004), Environmental assessment standards and risk management standards. http://www.astm.org/Standards/E2081.htm. Retrieved 10 Sept 2009

Athmer Ch (2004) In-situ remediation of TCE in clayey soils. Soil Sediment Contamination: Int J 13(5):381–390

ATSDR (2009) ToxFAQs^{TM} for total petroleum hydrocarbons (TPH) 1999. Agency for toxic substances and disease registry, http://www.atsdr.cdc.gov/tfacts123.html. Retrieved 10 Mar 2009

Balding DJ (2005) Weight-of-evidence for forensic DNA profiles. Wiley, West Sussex, England, ISBN 0470867647

Beard J, Australian Rural Health Research Collaboration (2005) DDT and human health. Sci Total Environ 355(1–3):78–89

Beck EC (1979) The love canal tragedy. EPA J. January 1979. http://www.epa.gov/history/topics/lovecanal/01.htm. Retrieved 30 Dec 2009

Braungart M, McDonough W (2002) Cradle to cradle. Remaking the way we make things. North Point, New York

Burmaster DE, Anderson PD (1994) Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. Risk Anal 14(4):477–481

Callahan B, Kostecki P, Reece K (2004) Human health risk assessment at a depleted uranium site. Soil Sediment Contam 13:597–609

Canadian Council of Ministers of the Environment (1999) Canadian environmental quality guidelines. Winnipeg (MB), Canada

Carlon C, Critto A, Marcomini A, Nathanail P (2000) Risk based characterisation of contaminated industrial site using multivariate and geostatistical tools. Environ Pollut 111(3): 417–427

Carlon C, Swartjes F (2007a) Analysis of variability and reasons of differences. In: Carlon C (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities. JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

Carlon C, Swartjes F (2007b) Rationale and methods of the review. In: Carlon C (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities. JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

Celico F, Musilli I, Naclerio G (2004) The impacts of pasture- and manure-spreading on microbial groundwater quality in carbonate aquifers. Environ Geol 46(2):233–236

Cobb GP, Sands K, Waters M, Wixson BG, Dorward-King E (2000) Accumulation of heavy metals by vegetables grown in mine wastes. Environ Toxicol Chem 19(3):600–607

Colvin VL (2003) The potential environmental impact of engineered nanomaterials. Nat Biotechnol 21:1166–1170

Commission of the European Communities (2006) Communication from the commission to the council, the European parliament, the European economic and social committee and the committee of the regions. Thematic Strategy for Soil Protection. COM(2006)231 final, Brussels, 22.9.2006

Crettaz P, Pennington D, Rhomberg L, Brand K, Jolliet O (2002) Assessing human health response in life cycle assessment using ED10s and DALY' s: part 1 – cancer effects. Risk Anal 22: 931–945

Darling CTR, Thomas VG (2003) The distribution of outdoor shooting ranges in Ontario and the potential for lead pollution of soil and water. Sci Total Environ 313(1–3):235–243

Dawson_JJC, Godsiffe EJ, Thompson IP, Ralebitso-Senior TK, Killham KS, Paton GI (2007) Application of biological indicators to assess recovery of hydrocarbon impacted soils. Soil Biol Biochem 39:164–177

DEFRA, EA (2002) DEFRA (Department for Environment, Food and Rural Affairs) and EA (The Environment Agency), EA R&D publication CLR GV 1–10, Bristol, UK

Duffus JH (2002) 'Heavy metals' a meaningless term? IUPAC Tech Rep Pure Appl Chem 74: 793–807

Edejer T, Baltussen R, Adam T (2003) Making choices in health care: WHO guide to cost-effectiveness analyses. World Health Organisation, Geneva, Switzerland

Eijsackers HJP, Groot M, Breure AM (2008) Upgrading system-oriented ecotoxicological research. Sci Total Environ 406(3):373–384. In: Breure AM, Eijsackers HJP, Groot M (eds) Ecological effects of diffuse pollution (Special issue)

European Commission (2009) http://ec.europa.eu/environment/soil/index_en.htm. Retrieved 8 Sept 2009

European Environmental Agency (2007) Overview of activities causing soil contamination in Europe. EEA, Copenhagen

European Union (2008) Directive 2004/42/CE of the European Parliament and of the Council of 21 April 2004 on the limitation of emissions of volatile organic compounds due to the use of organic solvents in certain paints and varnishes and vehicle refinishing products and amending Directive 1999/13/EC, EUR-Lex, European Union Publications Office. Retrieved on 17 Mar 2008

Fairhurst GT (1993) The leader-member exchange patterns of women leaders in industry: a discourse analysis. Commun Monogr 60:321–351

Ferguson C, Darmendrail D, Freier K, Jensen BK, Jensen J, Kasamas H, Urzelai A, Vegter J (1998) Risk assessment for contaminated sites in Europe. Volume 1. Scientific basis, Report of CARACAS project: concerted action on risk assessment for contaminated sites in the European Union. LQM, Nottingham

Fetzer JC (2000) Polycyclic aromatic hydrocarbons. Chemistry and analysis. A series of monographs on analytical chemistry and its applications, vol 158. Wiley-Interscience, New York

Filzek PDB, Spurgeon DJ, Broll G, Svendsen C, Hankard PK, Kammenga JE, Donker MH, Weeks JM (2004) Pedological characterisation of sites along a transect from a primary cadmium/lead/zinc smelting works. Ecotoxicology 13:725–737

Finlay PN (1994) Introducing decision support systems. NCC Blackwell; Blackwell, Oxford, UK; Cambridge, MA

Franken ROG, Baars AJ, Crommentuijn GH, Otte P (1999) A proposal for revised intervention values for petroleum hydrocarbons ('minerale olie') on base of fractions of petroleum hydrocarbons. RIVM report 711701015, RIVM, Bilthoven, the Netherlands

Friesl W, Horak O, Wenzel WW (2004) Immobilization of heavy metals in soils by the application of bauxite residues: pot experiments under field conditions. J Plant Nutr Soil Sci 167(1): 154–159

Friesl W, Lombi E, Horak O, Wenzel WW (2003) Immobilization of heavy metals in soils using inorganic amendments in a greenhouse study. J Plant Nutr Soil Sci 166(2):191–196

Grasmuck D, Scholz RW (2005) Risk perception of heavy metal soil contamination by high-exposed and low-exposed inhabitants: the role of knowledge and emotional concerns. Risk Anal 25:511–622

Grosse SD, Matte TD, Schwartz J, Jackson R (2002) Economic gains resulting from the reduction in children's exposure to lead in the U.S. Environ Health Persp 110:563–569

Guha B, Hills CD, Carey PJ, MacLeod CL (2006) Leaching of mercury from carbonated and non-carbonated cement-solidified dredged sediments. Soil Sediment Contamin: Int J 15(6):621–635

Hers I, Zapf-Gilje R, Johnson PC (2003) Evaluation of the Johnson and Ettinger model for prediction of indoor air quality. Ground water Monitor Remediation 23(2):119–133

Hjortenkrans D, Bergbäck B, Häggerud A (2006) New metal emission patterns in road traffic environments. Environ Monitor Assess 17(1–3):85–98

Hristov H, Butler N, Painter P, Siegel D (2005) Human-exposure-based screening numbers developed to aid estimation of cleanup costs for contaminated soil. Integrated risk assessment section office of environmental health hazard assessment. California Environmental Protection Agency, November 2004, January 2005 Revision

IPCS (2004) IPCS risk assessment terminology – harmonization project document no. 1. International programme on chemical safety. WHO, Geneva, Switzerland

Iturbe R, Flores C, Flores RM, Torres LG (2005) Subsoil TPH and other petroleum fractions-contamination levels in an oil storage and distribution station in north-central Mexico. Chemosphere 61(11):1618–1631

Jenck JF, Agterberg F, Droescher MJ (2004) Products and processes for a sustainable chemical industry: a review of achievements and prospects. Green Chem 6:544–556

Juhler RK, Felding G (2004) Monitoring methyl tertiary butyl ether (MTBE) and other organic micro pollutants in groundwater: results from the Danish national monitoring program. Water, Air, Soil Pollut 149(1–4):145–161

Kabir H, Raihan Taha M (2004) Assessment of physical properties of a granite residual soil as an isolation barrier. Elect J Geotech Eng 9 (Bundle B)

Kalis EJJ, Temminghoff EJM, Visser A, van Riemsdijk WH (2007) Metal uptake by Lolium perenne in contaminated soils using a four-step approach. Environ Toxicol Chem 26(2): 335–345

Kidd PS, Monterroso C (2005) Metal extraction by *Alyssum serpyllifolium* ssp. lusitanicum on mine-spoil soils from Spain. Sci Total Environ 336(1–3):1–11

Kjeldsen P (1999) Behaviour of cyanides in soil and groundwater: a review. Water, Air, Soil Pollut 115(1–4):279–308

Kremer RJ, Souissi Th (2001) Cyanide production by rhizobacteria and potential for suppression of weed seedling growth. Curr Microbiol 43:182–186

Lavado RS, Zubillaga MS, Alvarez R, Taboada MA (2004) Baseline levels of potentially toxic elements in pampas soils. Soil Sediment Contamin: Int J 13:329–339

Leijnse A, Hassanizadeh SM (1994) Model definition and model validation. Adv Water Resour 17:33:197–200

Levine AG (1982) Love canal: science, politics, and people. Lexington books, Lexington, MA, USA

Liu H, Probst A, Liao B (2005) Metal contamination of soils and crops affected by the Chenzhou lead/zinc mine spill (Hunan, China). Sci Total Environ 339(1–3):153–166

Lima ML (2004) On the influence of risk perception on mental health: living near an incinerator. J Environ Psychol 24(1):71–84

Lintelmann J, Katayama A, Kurihara N, Shore L, Wenzel A (2003) Endocrine disruptors in the environment, IUPAC Technical Report. Pure Appl Chem 75(5):631–681

Louekari K, Mroueh U-M, Maidell-Münster L, Valkonen S, Tuomi T, Savolainen K (2004) Reducing the risks of children living near the site of a former lead smeltery. Sci Total Environ 319(1–3):65–75

Macklin MG, Brewer PA, Balteanu D, Coulthard TJ, Driga B, Howard AJ, Zaharia S (2003) The long term fate and environmental significance of contaminant metals released by the January and March 2000 mining tailings dam failures in Maramure County, upper Tisa Basin, Romania. Appl Geochem 18(2):241–257

McHugh TE, Connor JA, Ahmad F, Newell ChJ (2003) A groundwater mass flux model for groundwater-to-indoor-air vapor intrusion, paper H-09. In: Magar VS, Kelley ME (eds) In situ and on-site bioremediation—2003. Proceedings of the seventh international in situ and on-site bioremediation symposium, Orlando, FL, June 2003

MERG (2001) Cyanide – the facts, MERG Report 2001–2, Mining Environment Research Group, by Laberge Environmental Services

Meuser H, Blume H-P (2001) Characteristics and classification of anthropogenic soils in the Osnabrück area, Germany. J Plant Nutr Soil Sci 164(4):351–358

Migliorini M, Pigino G, Bianchi N, Bernini F, Leonzio C (2004) The effects of heavy metal contamination on the soil arthropod community of a shooting range. Environ Pollut (2):331–340

Millis PR, Ramsey MH, John EA (2004) Heterogeneity of cadmium concentration in soil as a source of uncertainty in plant uptake and its implications for human health risk assessment. Sci Total Environ 326(1–3):49–53

Ministry of VROM (2008) Soil remediation circular 2006, as amended on 1 October 2008

Moon DH, Dermatas D, Menounou N (2004) Arsenic immobilization by calcium–arsenic precipitates in lime treated soils. Sci Total Environ 330(1–3):171–185

Nadal M, Schuhmacher M, Domingo JL (2004) Metal pollution of soils and vegetation in an area with petrochemical industry. Sci Total Environ 321(1–3):59–69

Nadim F, Hoag GE, Liu S, Carley RJ, Zack P (2000) Detection and remediation of soil and aquifer systems contaminated with petroleum products: an overview. J Petroleum Sci Eng 26(1–4):169–178

Nagy P, Bakonyi G, Bongers AMT, Kádár I, Fábián M, Kiss I (2004) Effects of microelements on soil nematode assemblages seven years after contaminating an agricultural field. Sci Total Environ 320(2–3):131–143

National Environmental Protection Council (2003) Guideline on the investigation levels for soil and groundwater, Canberra, Australia, National Environmental Protection Measure 1999, Schedule B(1)

NEN (2009) Soil – terrestrial soil – Strategy for the performance of the exploratory site investigation – Investigation of the environmental quality of soil and soil material (in Dutch), NEN (Dutch standardisation authority)

Norra S (2006) Urban soil science on the 18th WCSS, 18th world congress of soil science (WCSS), 9–15 July 2006, Philadelphia, PA. J Soils Sediments 6(3):189

Oliver L, Ferber U, Grimski D, Millar K, Nathanail P (2009) The scale and nature of European brownfields, CABERNET site. http://www.cabernet.org.uk/resourcefs/417.pdf. Retrieved 30 Dec 2009

Ollivon D, Blanchoud H, Motelay-Massei A, Garban B (2002) Atmospheric deposition of PAHs to an urban site, Paris, France. Atmos Environ 36(17):2891–2900

Oregon Department of Human Services (1994) Technical bulletin – health effects information. Department of Human Services Environmental toxicology section, Office of Environmental Public Health, February 1994, BTEX

O'Reilly M, Brink R (2006) Initial risk-based screening of potential brownfield development sites. Soil Sediment Contamin: Int J 15:463–470

Otte PF et al. (2009) Soil quality ambitions, route planner. http://www.soilambitions.eu. Retrieved on 19 Mar 2008

Pasquini MW, Alexander MJ (2004) Chemical properties of urban waste ash produced by open burning on the Jos plateau: implications for agriculture. Sci Total Environ 319(1–3):225–240

Pollard SJT, Lythgo M, Duarte-Davidson R (2002a) The extent of contaminated land problems and the scientific response. Issues Environ Sci Technol 16:1–19

Pollard SJ, Yearsley R, Reynard N, Meadowcroft IC, Duarte-Davidson R, Duerden SL (2002b) Current directions in the practice of environmental risk assessment in the United Kingdom. Environ Sci Technol 36(4):530–538

Prasad AS (1998) Zinc in human health: an update. J Trace Elements Exp Med 11(2–3):63–87

Provoost J, Cornelis Ch, Swartjes F (2006) Comparison of soil clean-up standards for trace elements between countries: why do they differ? J Soils Sediments 6(3):173–181

Qian X, Koerner RM, Gray DH (2002) Geotechnical aspects of landfill design and construction. Prentice Hall, Upper Saddle River, NJ

Raffensberger C, Tickner J (eds) (1999) Protecting public health and the environment: implementing the precautionary principle. Island, Washington, DC

Ramakrishnan U (2002) Prevalence of micronutrient malnutrition worldwide. Nutr Rev 60(Suppl 1):46–52

REVIT (2008) Working towards more effective and sustainable brownfield revitalization policies. Report of the Interreg IIIB report

Robinson BH, Mills TM, Petit D, Fung LE, Green SR, Clothier BE (2000) Natural and induced cadmium-accumulation in poplar and willow: implications for Phytoremediation. Plant Soil 227:301–306

Römkens FAM (1998) Effects of land use changes. On organic matter dynamics and trace metal solubility in soil. Dissertation University of Groningen, the Netherlands

Salemaa M, Derome J, Helmisaari H-S, Nieminen T, Vanha-Majamaa I (2004) Element accumulation in boreal bryophytes, lichens and vascular plants exposed to heavy metal and sulphur deposition in Finland. Sci Total Environ 324(1–3):141–160

Schuhmacher M, Granero S, Rivera J, Müller L, Llobet JM, Domingo JL (2000) Atmospheric deposition of PCDD/Fs near an old municipal solid waste incinerator: levels in soil and vegetation. Chemosphere 40(6):593–600

Seuntjes P (2004) Field-scale cadmium transport in a heterogeneous layered soil. Water, Air, Soil Pollut 140:401–423

Sivapullaiah PV, Lakshmikantha H (2004) Properties of fly ash as hydraulic barrier. Soil Sediment Contamin: Int J 13(5):391–406

Smith EP, Lipkovich I, Ye K (2002) Weight-of-evidence (WOE): quantitative estimation of probability of impairment for individual and multiple lines of evidence. Human Ecol Risk Assess 8(7):1585–1596

Souren AFMM (2006) Standards, soil, science and policy. Labelling usable knowledge for soil quality standards, PhD Thesis Free University Amsterdam, The Netherlands

Stockholm Convention (2009) Stockholm convention on persistent organic pollutants. About the convention. http://chm.pops.int/Convention/tabid/54/language/en-US/Default.aspx#convtext. Retrieved 30 Mar 2009

Suèr P, Nilsson-Påledal S, Norrman J (2004) LCA for site remediation: a literature review. Soil Sediment Contamin (formerly J Soil Contamin) 13(4):415–425

Swartjes FA (1999) Risk-based assessment of soil and groundwater quality in the Netherlands: standards and remediation urgency. Risk Anal 19(6):1235–1249

Swartjes FA (2007) Insight into the variation in calculated human exposure to soil contaminants using seven different European models. Integr Environ Assess Manage 3(3):322–332

Swartjes F, Carlon C (2007) Variability of soil screening values. In: Carlon C (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities, JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

Swartjes FA, d'Allesandro M, Cornelis Ch, Wcislo E, Müller D, Hazebrouck B, Jones C, Nathanail CP (2009) Towards consistency in risk assessment tools for contaminated site management in the EU. The HERACLES strategy from 2009 onwards. RIVM report 711701091, RIVM, Bilthoven, the Netherlands

Swartjes FA, Dirven-Van Breemen EM, Otte PF, Van Beelen P, Rikken MGJ, Tuinstra J, Spijker J, Lijzen JPA (2007) Towards a protocol for the site-specific human health risk assessment for consumption of vegetables from contaminated sites. RIVM report 711701040, RIVM, Bilthoven, the Netherlands

Swartjes FA, Tromp PC (2008) A tiered approach for the assessment of the human health risks of asbestos in soils. Soil Sediment Contamin 17:137–149

Suter GW II, Vermeire Th, Munns WR Jr, Sekizawa J (2005) An integrated framework for health and ecological risk assessment. Toxicol Appl Pharmacol 207:611–616

Sáková J, Tlustoš P, Pavlíková D, Hanč A, Batysta M (2007) Effect of addition of ameliorative materials on the distribution of As, Cd, Pb, and Zn in extractable soil fractions. Chem Pap 61(4):276–281

Teaf CM (1995) Human health and environmental risks associated with contaminated military sites. In: Herndon RC et al (eds) Clean-up of former Soviet Military installations. Springer, Berlin

Tlustoš P, Száková J, Kořínek K, Pavlíková D, Hanč A, Balík J (2006) The effect of liming on cadmium, lead, and zinc uptake reduction by spring wheat grown in contaminated soil. Plant Soil Environ 52(1):16–24

United Nations (1987) Report of the world commission on environment and development, A/RES/42/187, 96th plenary meeting, 11 December 1987

United Nations (1992) Report of the United Nations conference on environment and development, Rio de Janeiro, 3–14 June 1992, Annex I: Rio declaration on environment and development. A/CONF.151/26 (vol I)

US EPA (2002) Guidance for comparing background and chemical concentrations in soil for CERCLA sites. Office of emergency and remedial response U.S. environmental protection agency Washington, DC 20460. EPA 540-R-01–003 OSWER 9285.7–41 September 2002

US EPA (2007) Treatment technologies for site cleanup, 12th edn. US EPA Report 542R07012, September 2007

US EPA (2008) Title 40: protection of environment, part 51—requirements for preparation, adoption, and submittal of implementation plans. Subpart F—procedural requirements. § 51.100 Definitions. GPO Access: Electronic Code of Federal Regulations (e-CFR), U.S. Government Printing Office. Retrieved on 17 Mar 2008

US EPA (2009) US environmental protection agency website. http://www.epa.gov/brownfields/glossary.htm. Retrieved 30 Mar 2009

US National Research Council (1983) Risk assessment in the federal Government: managing the process. National Academy, Washington, DC

Vassilev A, Schwitzguebel JP, Thewys Th, Van der Lelie D, Van Gronsveld J (2004) The use of plants for remediation of metal-contaminated soils. Sci World 4:9–24

Von Lindern I, Spalinger S, Petroysan V, von Braun M (2003a) Assessing remedial effectiveness through the blood lead:soil/dust lead relationship at the Bunker Hill Superfund Site in the Silver Valley of Idaho. Sci Total Environ 303(1–2):139–170

Von Lindern IH, Spalinger SM, Bero BN, Petrosyan V, von Braun MC (2003b) The influence of soil remediation on lead in house dust. Sci Total Environ 303(1–2):59–78

Warren GP, Alloway BJ, Lepp NW, Singh B, Bochereau FJM, Penny C (2003) Field trials to assess the uptake of arsenic by vegetables from contaminated soils and soil remediation with iron oxides. Sci Total Environ 311:19–33

Weed DL (2005) Weight of evidence: a review of concept and methods. Risk Anal 25(6):1545–1557

Wessolek G (2006) Soil and art – the aesthetic of dirt, abstract for the 18th World congress of soil science (July 9–15, 2006), Philadelphia, PA. http://acs.confex.com/crops/wc2006/techprogram/P13610.HTM. Retrieved 3 Mar 2009

WHO (2001) Integrated risk assessment, report prepared for the WHO/UNEP/ILO international programme on chemical safety. WHO/IPCS/IRA/01/12. http://www.who.int/healthinfo/boddaly/en/. Retrieved 1 Aug 2008

WHO (2009) World health organisation website. http://www.who.int/healthinfo/boddaly/en/. Retrieved 30 Mar 2009

Williams ML, Locke VN (1999) Supervisor mentoring: does a female manager make a difference? Paper presented at the institute for behavioral and applied management (IBAM) conference. Annapolis, MD, November 4–6, 1999

Wong CSC, Xiang DL (2004) Pb contamination and isotopic composition of urban soils in Hong Kong. Sci Total Environ 319(1–3):185–195

Wong SC, Li XD, Zhang G, Qi SH, Min YS (2002) Heavy metals in agricultural soils of the Pearl River delta, South China. Environ Pollut 119(1):33–44

Yang QW, Shu WS, Qiu JW, Wang HB, Lan CY (2004) Lead in paddy soils and rice plants and its potential health risk around Lechang Lead/Zinc Mine, Guangdong, China. Environ Int 30(7):883–889

Yukselen MA, Alpaslan B (2001) Leaching of metals from soil contaminated by mining activities. J Hazardous Mater 87(1–3):289–300

Zhang G, Luo Y, Deng X (2002) Determination of surface area of red mud and beringite using methylene blue method. Pedosphere 12:189–192

Zytner RG, Salb AC, Stiver WH (2006) Bioremediation of diesel fuel contaminated soil: comparison of individual compounds to complex mixtures. Soil Sediment Contamin: Int J 15(3):277–297

Chapter 2
Characteristics of Natural and Urban Soils

Helmut Meuser and Robert H.M. Van de Graaff

Abstract This chapter deals with soils that have been contaminated by human activities and soils that have inherent contamination due to natural causes, as well as soils that may be thought of as contaminated, but where the contaminants are part of the natural geochemistry, they are completely inert and unlikely to cause any significant risk to life. Contaminated soils have figured greatly in the mind of the public since a number of dramatic, highly publicised events occurred in which residential developments on former landfills caused serious health problems for the inhabitants. Logically, the primary focus of studies and inventories of soil contamination were the industrial lands, from medieval metal processing to modern day manufacturing and storage, war zones and battle grounds. Besides, other factors impacting contamination potential have to be taken into account, e.g. flood occurrences and ubiquitous atmospheric deposition. There was a need to define background levels of all contaminants that can occur naturally in the land or could have been added by humans. Likewise, a need to discover the mobility and toxicity of the contaminants was required to develop Risk Assessment. Finally, chemical affinities and solubilities of several contaminant groups are discussed.

Contents

H. Meuser (✉)
University of Applied Sciences, Osnabrück, Germany
e-mail: h.meuser@fh-osnabrueck.de

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_2, © Springer Science+Business Media B.V. 2011

2.1 Soils of Contaminated Sites

This chapter deals with soils that have been contaminated by human activities and soils that have inherent contamination due to natural causes, as well as soils that may be thought of as contaminated, but where the contaminants are part of the natural geochemistry comprehensive data about contaminated urban soils can be found in Meuser 2010.

2.1.1 Natural and Anthropogenic Soils

Soil contamination can relate to all kind of soil materials and the subsequent distribution of contaminants within the soil profile is often controlled by the build-up and composition of soil layers and horizons. In general, we have to differentiate between three groups of soils that may have become contaminated:

1. Soils derived from natural parent material, indicating natural pedogenesis, such as Cambisols, Gleysols, Lixisols, Planosols, Solonchaks, Ferralsols and many other Major Soil Groupings (FAO 2006; FAO-Unesco 1998). The differentiation between all these Major Soil Groups is based on transformations caused by their weathering histories and differences between their original parent materials. Also, the movement and accumulation of contaminants in these natural soil profiles is affected by the horizonation of the profile and the varying permeability of individual soil horizons.
2. Natural soils under mainly long-term horticultural cultivation (Anthrosols; FAO 2006), which commonly are further specified in, for instance, the German nomenclature as Hortisols and Treposols (Meuser and Blume 2004). These soils are the most likely soils to have been treated with high amounts of nutrients, especially P and N, from farmyard manures and fertilisers, as well as metals

such as arsenic and copper (from fungicides), and organic contaminants such as DDT and dieldrin (from pesticides).

3. Urban soils from an anthropogenic origin (Technosols, FAO 2006), usually deposited, consisting of technogenic man-made material or of both technogenic and natural material (Deposols in German nomenclature). These soils are very likely to contain contaminants, as they are often waste soils from excavations at urban sites and can contain almost anything, from rubble to mine waste and fly ash.

The second and third group belong to the anthropogenic soils. The FAO-Unesco (1998) soil classification system World Reference Base (WRB) refers to the second group as Anthrosols (soils with long and intensive agricultural use), and the third one is called Technosols (soils containing many artefacts).

The second group, associated with long-term horticultural or agricultural use, indicates man-made profile changes. The soils, however, are based on natural parent material. Typically, the humic topsoil is deeper than the usually ploughed soil depth of 30 cm. We can distinguish between several kinds of soils within this group (Meuser and Blume 2004):

- *Garden soil*, significant in residential areas with gardens surrounding detached houses, row houses, or exclusive residences, as well as for allotments. Continuous organic matter application (e.g., compost, manure), combined with frequent digging procedures, leads to humic top soils with an enhanced biological activity. Based upon the German nomenclature, garden soils with a humic topsoil of at least 40 cm are termed *Hortisols*. Different from garden soils, there are soils of cemeteries and churchyards, indicative of deep digging activities to a depth of 180 cm and ultimately humus accumulation.
- Moreover, there are so-called *plaggen soils* occupying large areas of the northern parts of Germany and the Netherlands. They result from long-term agricultural use in time periods between the 14th and 19th century, when farmers used organic manure only, leading to a raised soil surface in the course of time. They are of importance at urban sites as well, since they are mostly located close to built-up areas, in particular in the peri-urban area.

Generally, the cultivated soils of concern show enhanced contaminant concentration due to atmospheric input (air pollution) and artificial amendments like sewage sludge or ashes. For instance, many urban garden soils have been fertilized with ashes derived from domestic home heating processes, leading to a high concentration of heavy metals compared with neighbouring cropland sites (Meuser and Anlauf 2007) (Fig. 2.1).

Anthropogenic soils, or Deposols, can be subdivided according to distinct *kinds of deposits* (Fig. 2.2). In the landscape, deposits consisting of mining heaps, waste landfills, dumps, and linear structures such as dikes, dams, and sound-isolating earth walls, are easily recognisable. The fill materials, however, result from excavations

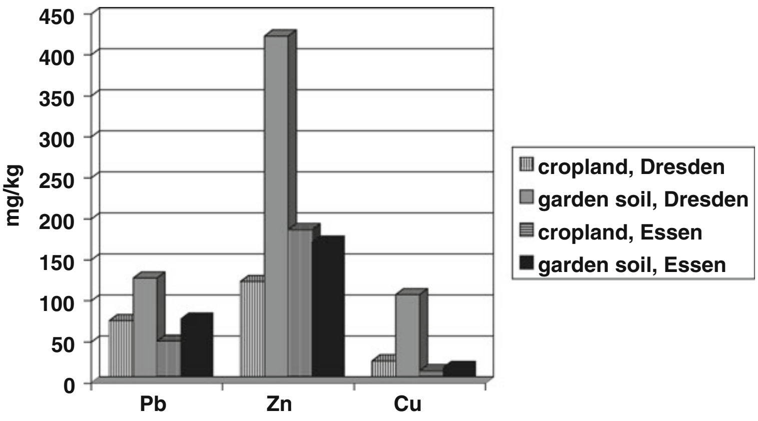


Fig. 2.1 Pb, Zn and Cu content (mg kg^{-1}) of garden soils compared with neighbouring cropland sites in Essen and Dresden, Germany (Meuser and Anlauf 2007)

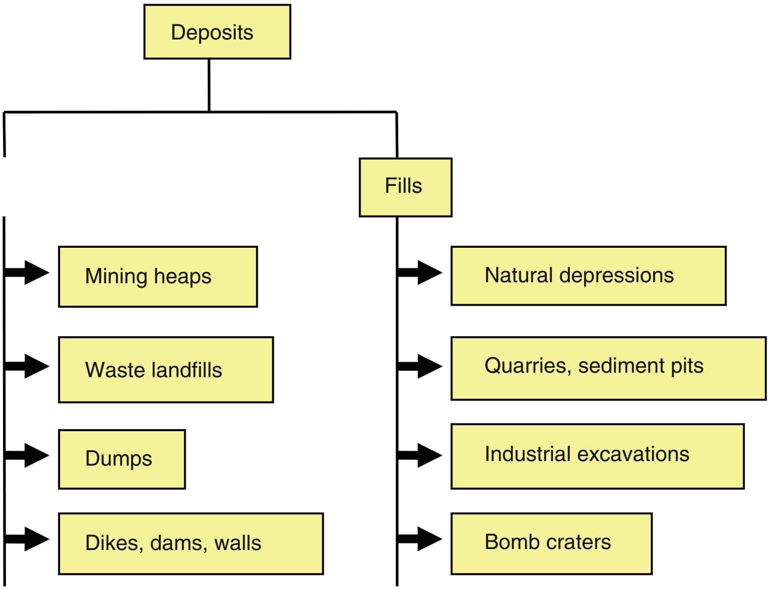


Fig. 2.2 Overview on distinct kinds of deposits

or refilling of former depressions. Consequently, they do not have to be recognisably similar to the materials in their natural source areas. Fills are associated with former usually watery natural depressions (swamp, peat land), quarries (sandstone, limestone et cetera), and pits (sand, clay, loam, marble), leading to a general raising of the natural ground level (Nathanail and Bardos 2004). Occasionally, they are located in industrial areas, where excavation has taken place, or they are the result from warfare such as bomb craters.

During the Balkan War in 1999, strategically important targets were destroyed, but also collateral damage (e.g., treatment facilities for waste and wastewater, fertilizer plants) was of importance. Regardless of the nature of the targets, a release of hazardous substances into the environment, particularly into the soil, had to be expected. Furthermore, attention must be paid to unexploded bombs and landmines indicating a serious redevelopment problem. In agricultural land, a rapid refill of bomb craters with distinct and partly unknown materials had to take place in order to grow staple food crops. A single 500 pound explosive bomb is able to cause a bomb crater up to 14 m wide and 9 m deep. No one knows the number of bomb craters in European countries resulting from the Second World War, but most of the craters were refilled rapidly without any knowledge about the materials that were used. The number of these craters could be high. For instance, it is assumed that in the Indochina conflict 25 million craters were caused by air raids (Genske 2003).

The frequency of technogenic materials occurring in deposited soils (termed *Deposols*) is variable, but some general principles are possible to report. The frequency decreases with increasing distance from the urban agglomerations, in particular city centres. Nevertheless, even in the countryside former excavated areas had been refilled with anthropogenic materials in order to level the soil surface or to improve the physical soil quality in wetlands, for construction purposes. Therefore, deposited soils that contain transported technogenic materials from other sites far away can be present in almost every area.

Considering the soil inventory of the German city of Muenster (Fig. 2.3), soils in the city centre as well as in industrially and administratively used areas reveal a lot of deposited materials, while garden soils dominate the residential areas, and particularly allotments. Furthermore, soils from deposited material can be found in parks (refilling processes after the Second World War) and sports fields consisting of man-made, compacted soils to enable sports activities (Meuser 2007). Craul (1992) reported that in the U.S. capital Washington D.C., located in a coastal area, approximately 81% of the area is disturbed and 14% must be classified as fills. A park soil survey in the city of Hamburg, Germany, showed 26.5% deposited soils, 26.5% containing natural material and an additional 20.5% with a mixture of both natural and technogenic material (Däumling 2000). In the city of Stuttgart, Germany, 53% of the upper 100 cm of the systematically investigated soil revealed more or less technogenic ingredients (Stasch et al. 2000).

In parks and lawn areas that surround houses, often beautifully landscaped with flowers and ornamental shrubs, deposited soils may only be covered by a thin humic topsoil layer. In parks of Washington D.C., the thickness of the humic top soils range between 6 and 35 cm (Craul 1992). It is easier to understand that contaminated sites are associated with derelict and abandoned areas. These areas are normally recognisable by visible ecological and social impacts such as unvegetated soils, subsidence, ruined buildings, absence of infrastructure, unsightly neighbourhood, and poor housing, respectively.

The example of man-made, compacted soils on the sports fields, mentioned above, demonstrates the different soil functions found in urban areas. Apart from the natural functions of soil, such as providing a medium for plant growth, a recycling

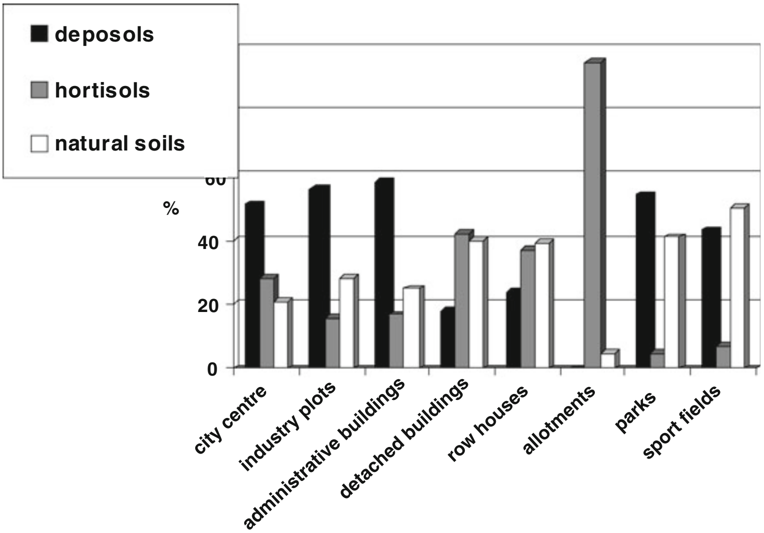


Fig. 2.3 Percentage of deposols, garden soils (Hortisols based on German nomenclature) and natural soils according to the use in Muenster, Germany (Meuser 2007)

system for nutrients, a system for water supply and purification, and a habitat for soil organisms, soils serve as an engineering medium on which to base such structures as roadbeds, building foundations, and sport parks (Brady and Weil 2008).

The land use type may also be decisive with respect to soil contamination. Biasioli et al. (2008) mentioned that the concentrations of the typical anthropogenically impacted metals, copper, lead, and zinc, exhibited their highest values in ornamental gardens and at roadsides in the city centre and residential areas, followed by open spaces (parks) and ultimately riverbanks. This phenomenon was found through analysis of soils in completely different European cities (Ljubljana, Slovenia; Seville, Spain; Torino, Italy).

2.1.2 Imported Filling Materials

Soils made up of imported materials are frequently contaminated irrespective of the kind of material deposited. Independent of the reasons for mound construction and refilling processes it should be taken into consideration that not only natural materials have been used. To the contrary, waste products including potentially toxic substances are generally utilised for mounds and fills, possibly for disposal of these unwanted materials. In Table 2.1 several technogenic components are listed that can be expected in heaps and fills (Meuser 1996). Technogenic material is completely artificially produced (e.g., slag), or at least anthropogenically changed or manipulated, like coal gangue.

Table 2.1 Material compositions of heaps and fills

Main component group	Component group	Examples
Construction rubble	Housing/industrial debris	Brick, mortar, concrete, plaster
	Asphalt debris	Bitumen and tar asphalt
Slag	Iron works	Blast furnace slag, sand and pumice of iron works
	Steel works	Slag of steel smelting furnace
	Heavy metal works	Lead slag, copper slag, zinc slag
	Foundry	Slag and sand of foundry
Ashes	Hard coal-fired and lignite coal-fired power station	Fly ash, bottom ash
	Garbage incinerator	Fly ash, bottom ash
Mining waste	Coal mining waste	Coal gangue
	Ore mining waste	Ore gangue
Refuse	Household refuse	Glass, metal, paper, ceramics, organic garbage, wood, bulky refuse
	Industrial refuse	Dross, cinders, salt slag, other specific waste of industrial processes
Sludge	Sewage and wastewater works	Sewage sludge
	Harbor and river dredging	Harbor and river sludge
	Solution mining	Ore and salt mining sludge

Meuser (1996)

The use of such materials as mono-substrates can be supposed to be relatively rare. In most of the cases, mixtures of technogenic and natural materials are typical for deposited soils. Therefore, deposited soils with a high material heterogeneity and lithologic discontinuity, which reveal great vertical and spatial variability, make survey and sampling difficult.

2.2 Inherited Geochemistry

This section deals with abundances and the sometimes elevated concentrations of trace elements in natural materials (Krauskopf 1967; Reimann and De Caritat 1998) that may be erroneously considered to be contaminated sites. In the great majority of cases where "unusually" high concentrations of heavy metals have been found, their solubility and mobility are extremely low so that ecological impacts are negligible. In some cases, however, such as the high nickel concentration in soils formed on serpentine rock, the vegetation has evolved to cope with the higher availability of nickel. Reimann and de Caritat (1998) warn that locally elevated concentrations of a specified element in a rock type will have an influence on all media that interact with that rock, such as soils, groundwater, vegetation and wind blown dust. Likewise their readers are told to treat so-called “average” data on concentrations for specific media with care as the sampling and analytical methods are often unknown.

Metals and metalloids and other trace elements vary greatly in natural abundance in different geological materials (Krauskopf 1967; Reimann and De Caritat 1998), as shown in Table 2.2.

Basalt-derived soils in low rainfall climates that have not undergone much leaching invariably are high in calcium (Ca) and magnesium (Mg) as well as in several metals, such as cobalt, chromium, copper, iron, manganese, nickel, and vanadium. Other rock types such as granites or shales have substantially lower metal concentrations, but in the soils that formed on these rocks these metals often have been translocated to the subsoil. High arsenic is common in soils on the edges of certain mineralised ore bodies. Another example of naturally elevated metal concentrations is the high zinc soils in parts of the Netherlands and Belgium. Their abundance in soils derived from these lithologies is determined by the abundance in the parent material, but strongly affected by subsequent weathering and geochemical processes, as well as by the mixing that occurs when unconsolidated earth materials undergo transportation by wind, water or ice sheets.

Weathering processes operating on fresh rocks as well as previously weathered geological materials will transform and release many or most of these elements, with secondary minerals forming and capturing particular trace elements in their lattices, according to their geochemical affinities. In high rainfall areas, well weathered basalt soils have undergone high losses of silica, calcium and magnesium, but the heavy metal concentrations have remained high due to their transformation to low-solubility oxides/hydroxides.

Large areas of the world are covered in geologically recent (Holocene) alluvial or windblown sediments or glacial morainic materials. Loess and morainic deposits are widespread in Northern North America, Northern Europe and Northern Asia. These consist of transported soils, having often a mixed geological origin and having undergone relatively little weathering, as they tend to be of Post-Glacial age and, hence, are younger than 12,000–16,000 years. All the great mountain belts and hilly areas of the world are affected by continuing erosion exposing freshly weathered soil material on slopes and colluvial mantles.

Elsewhere, large areas of the world, particularly in the continents of Africa, South America and Australia, consist of old Tertiary-age land surfaces covered by a deep regolith that is largely in situ. Regolith[1] is a general term for the layer of unconsolidated (non-cemented), weathered material, including rock fragments, mineral grains and all other superficial deposits, that rests on unaltered, solid bedrock. Its lower limit is the weathering front. Soil is regolith that often contains organic material

---

[1] Oxford Dictionary of Earth Sciences, Oxford University Press, 1999.

Table 2.2 Average abundance of major and trace elements in major rock types and seawater in parts per million (ppm)

Element	Continental crusts	Granite	Basalt	Shale	Seawater
O	46.4×10^4.	–	–	–	857,000
Si	28.2×10^4	32.3×10^4	24×10^4	23.8×10^4	3.0
Al	8.2×10^4	7.7×10^4	8.8×10^4	8.0×10^4	0.01
Fe	5.6×10^4	2.7×10^4	8.6×10^4	4.7×10^4	0.01
Ca	4.1×10^4	1.6×10^4	6.7×10^4	2.5×10^4	400
Na	$2.4\ 10^4$	2.8×10^4	1.9×10^4	0.66×10^4	10,500
Mg	2.3×10^4	0.16×10^4	4.5×10^4	1.34×10^4	1,350
K	2.1×10^4	3.3×10^4	0.83×10^4	2.3×10^4	380
Ti	5,700	2,300	9,000	4,500	0.001
H	1,400	–	–	–	108,000
P	1,050	700	1,400	770	0.07
Mn	950	400	1,500	850	0.002
F	625	850	400	500	1.3
Ba	425	600	250	580	0.03
Sr	375	285	465	450	8.0
S	260	270	250	220	885
C	200	300	100	1,000	28
Zr	165	180	150	200	–
V	135	20	250	130	0.002
Cl	130	200	60	160	19,000
Cr	100	4	200	100	0.000,05
Rb	90	150	30	140	0.12
Ni	75	0.5	150	95	0.002
Zn	70	40	100	80	0.01
Ce	67	87	48	50	5.2×10^{-6}
Cu	55	10	100	57	0.003
Y	33	40	25	30	0.000,3
Nd	28	35	20	23	9.2×10^{-6}
La	25	40	10	40	1.2×10^{-5}
Co	25	1	48	20	0.000,1
Sc	22	5	38	10	0.000,04
Li	20	30	10	60	0.17
N	20	20	20	60	0.5
Nb	20	20	20	20	0.000,01
Ga	15	18	12	19	0.000,03
Pb	12.5	20	5	20	0.000,03
B	10	15	5	100	4.6
Th	9.6	17	2.2	11	0.000,05
Sm	7.3	9.4	5.3	6.5	1.7×10^{-6}
Gd	7.3	9.4	5.3	6.5	2.4×10^{-6}
Pr	6.5	8.3	4.6	5	2.6×10^{-6}
Dy	5.2	6.7	3.8	4.5	2.9×10^{-6}
Yb	3	3.8	2.1	3	2.0×10^{-6}
Hf	3	4	2	6	–
Cs	3	5	1	5	0.000,5
Be	2.8	5	0.5	3	6×10^{-7}
Er	2.8	3.8	2.1	2.5	2.4×10^{-6}

Table 2.2 (continued)

Element	Continental crusts	Granite	Basalt	Shale	Seawater
U	2.7	4.8	0.6	3.2	0.003
Br	2.5	1.3	3.6	6	65
Sn	2	3	1	6	0.000,8
As	1.8	1.5	2	6.6	0.003
Ge	1.5	1.5	1.5	2	0.000,06
Mo	1.5	2	1	2	0.01
W	1.5	2	1	2	0.000,1
Ho	1.5	1.9	1.1	1	8.8×10^{-7}
Eu	1.2	1.5	0.8	1	4.6×10^{-7}
Tb	1.1	1.5	0.8	0.9	–
Lu	0.8	1.1	0.6	0.7	4.8×10^{-7}
Tm	0.25	0.3	0.2	0.25	5.2×10^{-7}
I	0.5	0.5	0.5	1	0.06
Tl	0.45	0.75	0.1	1	<0.000,01
Cd	0.2	0.2	0.2	0.3	0.000,11
Sb	0.2	0.2	0.2	1.5	0.000,5
Bi	0.17	0.18	0.15	0.01	0.000,02
In	0.1	0.1	0.1	0.05	<0.02
Hg	0.08	0.08	0.08	0.4	0.000,03
Ag	0.07	0.04	0.1	0.1	0.000,04
Se	0.05	0.05	0.05	0.06	0.000,4

Au, Pt metals, Re and Te are less than 0.05 ppm in rocks and less than 0.000,01 ppm in seawater. Inert gases He, N, Ar, Kr and Xe are not listed as they only occur in seawater at low concentrations Krauskopf (1967)

and is able to support rooted plants. The lower part of a weathering profile is called saprolite[2] or the chemically rotted rock in situ.

Depending on the climatic regime that exists in a region where rocks or soils continue to weather, the chemical transformations of mineral parent material are affected by the complete or partial removal, or merely local translocation, of soluble minerals and by the pH of the system. Rocks and soils consist largely of silicate minerals. Their transformations to secondary minerals and solutes are reasonably well known and documented. McBride (1994) summarises these weathering environments as follows:

- *Strongly alkaline, confined environments*: Smectite clays and other silica-rich silicate minerals are formed and potassium, calcium and magnesium salts may accumulate.
- *Moderately alkaline, weak-leaching environments*: The most soluble by-products of weathering are removed, such as the base metals and some of the silica, but secondary clays as smectites and illites remain.

---

[2] Ibid.

- *Acidic, strong leaching environment*: The ultimate end products formed after prolonged weathering include kaolinite clay, alumina (bauxite) and iron hydroxides and oxides.

Weathering, leaching, pH and internal (within soil profile) translocation in all three environments affect also the distribution and mobility of those chemical elements such as cadmium and lead that are recognised as contaminants, regardless of whether they were native or have been added by mankind.

Most of Northern and Western Europe are in the third class of environments (*acidic, strong leaching environment)*. But they still have youthful landscapes and soils, as they are post-Glacial and rainfall is sufficient to bring about leaching and soil pH will be lower than 6, unless the parent lithology is calcareous or the soils have been amended by concrete, mortar or other technogenic materials. Southern Europe is characterised by weak leaching environments, and, where erosion and sedimentation is active, by youthful landscapes also. The environments of China, Canada and much of the United States of America occupy all three classes, and on the whole may also be interpreted as youthful landscapes due to active erosion.

The Australian continent has examples of weathering environments that originally were acidic and strongly leaching, but due to their ancient origin and changed climate are now confined or weakly leaching, thus preserving characteristics that belong to the distant past. It also contains strongly alkaline confined environments.

On a much more detailed scale, the Holocene soils of the Netherlands, which are largely alluvial and glacial terminal morainic sediments and peat deposits, have been subdivided into four broad soil categories in order to understand the characteristic trace element background concentrations. They have been found to display specific geochemical affinities.

Figure 2.4 showing the generalised soil map of the Netherlands may be compared to the generalised soil map of the area around Melbourne, Victoria, Australia, in Fig. 2.5.

The soils around Melbourne are mapped on the basis of underlying geological parent materials, with an overlay of weathering history and soil development, *i.e.*, soil age, and represent local regoliths. Thus, the red loams and the brown loams over clay are ancient well-leached, acidic deep soils. They were formed on Pliocene-Eocene basalts during a former wet and warm climate, and where the present climatic regime produces some 850 mm yr^{-1} rainfall and 1,060 mm yr^{-1} evaporation, with the main rainfall occurring during the colder winter months. Thus the weathering and soil formation goes back to the early Pleistocene at the least.

In the far west of the Melbourne map, the heavy clays on Quaternary basalt all have strongly alkaline and sodic sub-soils with lime accumulations at 0.5 m or deeper, because there can be no significant leaching with an annual evaporation of more than 1,600 mm and annual rainfall of 550 mm.

Deep sands free of lime are Pleistocene windblown highly leached and acidic sand sheets with podzol profiles, while the deep sands with lime are young Holocene calcareous dunes, whose carbonate content derives from seashells.

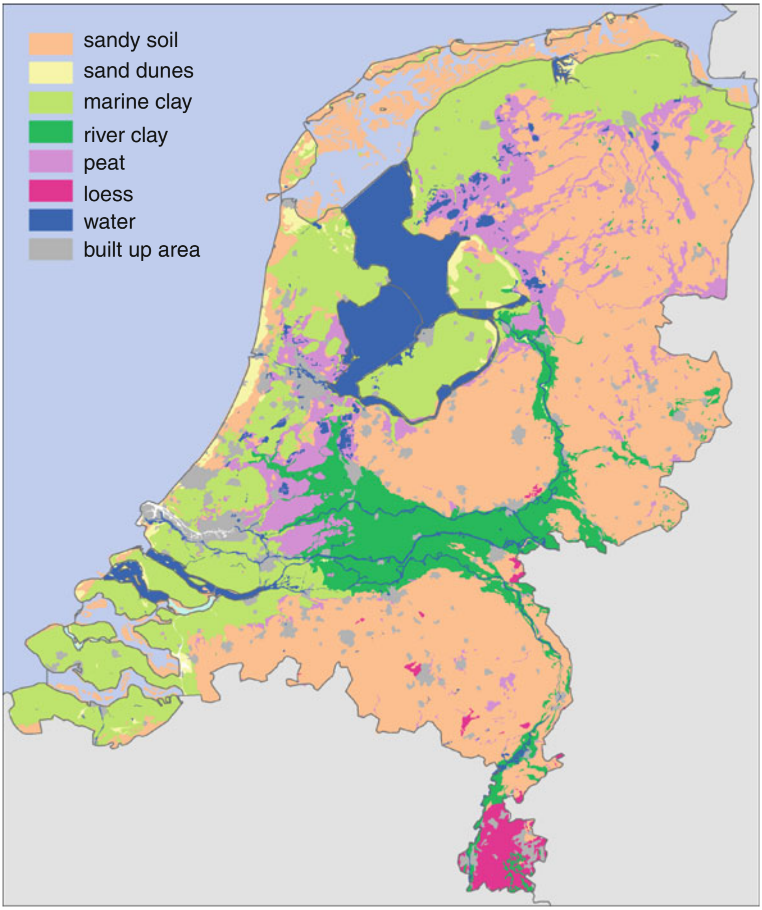


Fig. 2.4 Generalised map of dominant soil types in The Netherlands, as an example of a soil map, based on main soil texture classes (Fraters et al. 2001)

Light grey loams over clay are profiles having generally abrupt textural increases from A to B horizon in profiles formed from mainly Silurian interbedded shales, mudstones and sandstones and are probably older than Holocene. They are more acidic in the high rainfall eastern zone than in the lower rainfall central northern zone. These major mapping units have distinct geochemical assemblages in terms of heavy metals.

Soil contamination generally is a superimposed change of the natural soils. Over and above the natural geochemical characteristics of soils, the Old World has received anthropogenic additions of trace elements, as has been the case ever since ores were first smelted to make metal atmospheric deposition of metals. Even in the Greenland ice cap there are traces of lead derived from the Mediterranean smelters. Industrialisation took off on a grand scale more than two centuries ago and contributed more atmospheric deposition of inorganic contaminants. Agricultural

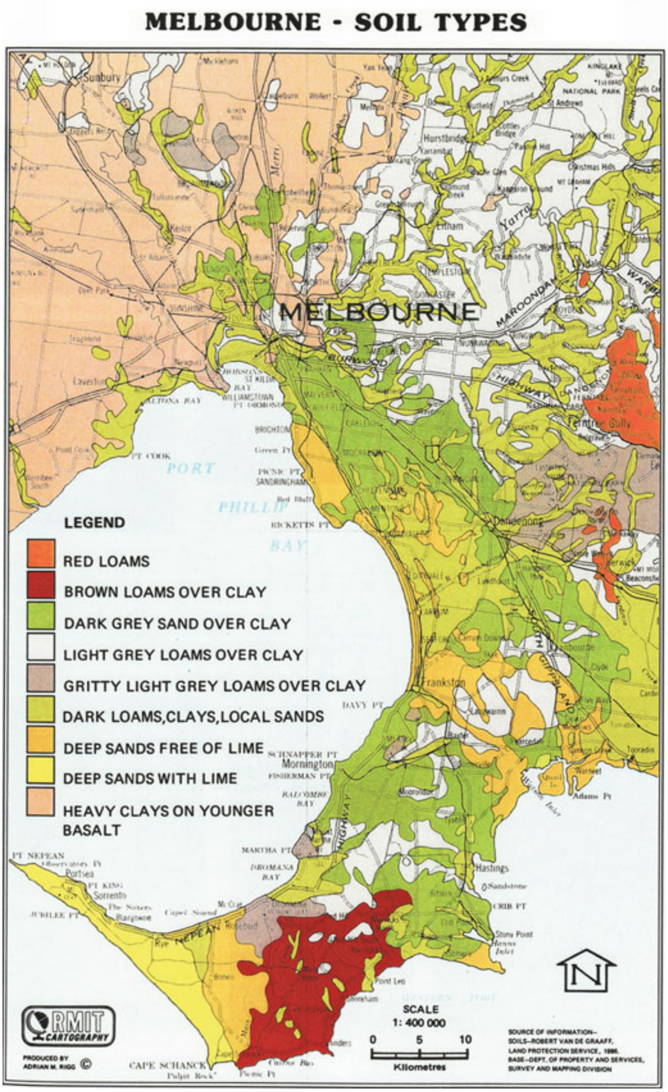


Fig. 2.5 Main broad, geologically-related soil types of the Melbourne Metropolitan area, as an example, and surroundings. The East-West width of the map is approximately 60 km. From west to east there is a rainfall gradient from about 550 mm yr^{-1} to more than 1,150 mm yr^{-1}, and evaporation from about 1,600 mm yr^{-1} to 1,100 mm yr^{-1} (details not clearly visible; both graphs are meant to give a general insight into soil maps)

use of copper, arsenic and lead in pesticides has also been intensive. In contrast, the Australian continent has been too distant from these industrial sources to have been significantly affected and Australia's own industrial and agricultural development has made much less of an imprint in terms of these inorganic contaminants. However close to the main urban centres anthropogenic contamination is superimposed on the natural system.

Figure 2.6 illustrates the ranges of trace elements that can be found in mineral soils. It is probably based on data obtained in North America, and takes no account of the particular soil (weathering) environments from which the samples were collected.

In many countries methods and criteria for assessing impacts on human health, ecology and groundwater have developed. Thus, Soil Quality Standards based mainly on total concentration were carried out.

It should also be noted that the entire focus is always on the known contaminant metals or metalloids and takes no account of other elements or soil properties that play a major role in immobilising the contaminant elements. Likewise, the role of substances like chloride, which can render metals such as lead much more soluble and mobile by complexation, or soil pH which also affects speciation and mobility, is ignored in these criteria. All soils are, as it were, regarded as being a "standard soil" in Dutch practice. The exclusive focus on those elements that are known or suspected to be of public health and ecological significance resulted in establishment of companies that carry out contamination assessments and site remediation, determining only those elements in laboratory analyses. Therefore intelligent assessment is deprived of relevant information.

Figure 2.6 was taken from Mitchell (1965). The stippled sections indicate more unusual concentration values. Certain extremely high values from localities

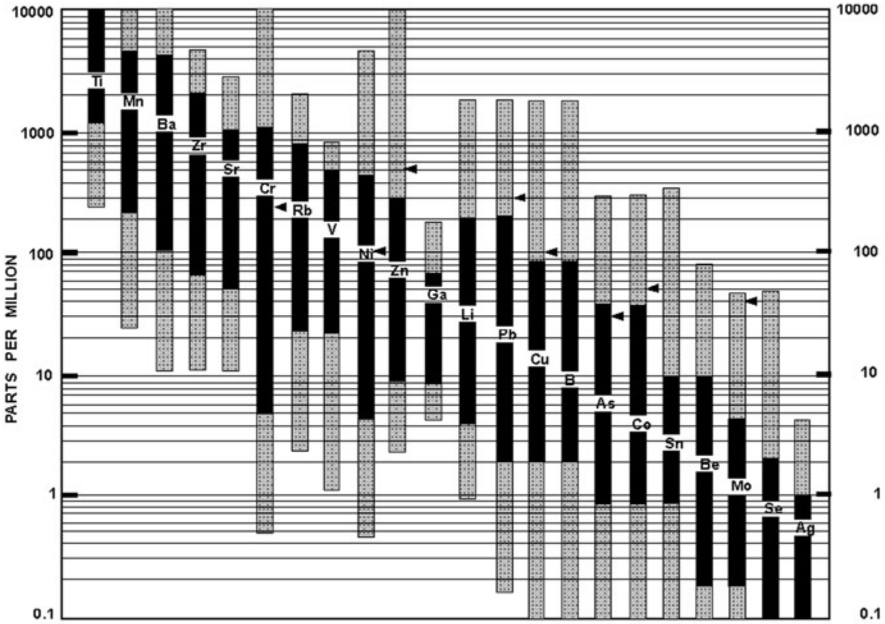


Fig. 2.6 The ranges of trace elements that can be found in mineral soils (Mitchell 1965)

influenced by ore deposits have been ignored. The arrows represent maximum limits for clean fill set by the EPAV.[3] Figure 2.6 clearly shows that "background" concentrations vary enormously, because the data represent a very large range of soils of different derivations without necessarily the ecology dependent on these soils showing any toxicity symptoms. Therefore total concentrations *per se* have no informative content with regard to ecological toxicity. Hodson et al. (Chapter 16 of this book) give a detailed analyses of bioavailability.

In layered and faulted geological materials it is common to find that preferred pathways for groundwater seepage have created zones of high accumulation of iron and associated trace metals and metalloids rendering the concept of natural background concentrations unworkable. These fracture zones often develop high concentrations of ferruginous material. Similarly, the process of vertical leaching of rainwater through a soil profile, especially if it contains layers of contrasting permeability, will distribute trace elements in a non-uniform manner through the profile.

Moreover, in Australia, the sharp boundary between a lighter-textured topsoil and a clay subsoil, in soil profiles that are commonly called Duplex profiles (Northcote 1960), often becomes a localised periodic low redox/high redox zone where ferruginous gravels form by accretion of iron oxides/hydroxides on existing nuclei. These become sites of preference for the accumulation of arsenic, vanadium, nickel, chromium, cobalt and other metals if present in the material. For other kinds of soil profiles one often finds that native inorganic contaminant concentrations form a depth function similar to a chromatogram, in much the same way that calcium carbonate and gypsum accumulate at a preferred depth in the soil profile.

Clearly, for these regolith bodies, which can have such complicated contaminant distributions, it is not logical to maintain individual background concentrations for trace elements to be used in soil related policy. However, so long as environmental regulators in Australia continue to rely on single background concentration standards to determine where further investigations and assessments are required, then the risk assessors must be free to apply the guidelines intelligently and the regulators must be open to scientific argument in individual cases.

2.3 Contaminant Behaviour in Soils

2.3.1 Chemical Affinities and Solubilities

Contaminants occur in many soils as a result of having been added to the soil by mankind or due to inherited geological processes. An understanding of the origin and mobility of contaminants is essential in making judgements on their human health and ecological impacts. In this section the characteristics of contaminant groups is described in terms of affinities and solubilities.

---

[3] Environment Protection Authority, Victoria, Australia.

2.3.1.1 Acids and Bases

Acids and bases all are capable of dissociating in water and affecting the pH of the soil. They are usually highly soluble in water, even if they are weak acids or bases, and therefore they are mobile. By affecting the pH of the soil, they can change the solubility and mobility of contaminants, particularly heavy metals and metalloids. Soils with a significant cation exchange capacity (CEC) can adsorb a high amount of the proton, H^+, and other positively charged molecules. By the same token, high CEC soils can buffer external acidifying and alkalifying input better.

Liquid acidic chemical wastes that may have been applied to the soil include diluted HCl, H_2SO_4, H_3PO_4 and organic acids such as acetic acid. Alkaline liquid wastes include Na(OH) solutions, other caustic solutions, hydrazobenzene solutions, or solid phase burnt lime CaO. These compounds can have large effects on the permeability of the soil and its shrinkage properties, as well as cause direct dissolution of the soil (Dragun 2007). Depending on the soil type, extreme changes in soil pH also can destroy living organisms.

Soils may become very acid, anthropogenically, by long-term and copious use of certain fertilizers like ammonium nitrate, ammonium sulphate, and monocalcium phosphate (MCP) as autotrophic bacteria convert the ammonium ion to $2H^+$ and $NO_3{}^-$, and as the MCP hydrolyses to form dicalcium phosphate and phosphoric acid. Mine wastes or dredged soils that include sulphide minerals, such as pyrite, FeS_2, can also create extreme acidity through oxidation of the S to $SO_4{}^{2-}$. This process is called "acid sulphate soil formation" and can lead to extremely low pH values that are lethal to plants and soil organisms and can cause acid runoff that might kill fish and other aquatic life forms. Such an oxidation process was already identified in the eighteenth century by the Swedish botanist Linnaeus (1707–1778). However, soil minerals such as carbonates, alumino-silicate clays and oxide and hydroxide minerals can act as buffers to excessive acidity by reacting with the H^+ ions, dissolving and forming water and very weak acids.

Soils affected by extreme acidity can be neutralised by a variety of liming materials such as calcium oxide, calcium hydroxide, calcium carbonate, and dolomite. Extreme alkalinity can be neutralised by adding elemental sulphur, aluminium sulphate, or calcium polysulphide.

2.3.1.2 Water-Immiscible Contaminants

Water-immiscible contaminants include petroleum hydrocarbons such as NAPL, LNAPL, DNAPL. In an extensive discussion of this topic, Dragun (2007) states that the molecular structure of these contaminants is the chief determinant of the strength of adsorption as structure determines the molecular volume, often also the water solubility, the octanol-water partition coefficient, and the vapour pressure. The larger the molecule, the smaller will be its polarity, and the more important will be the van der Waals forces' attraction to the solid phase surfaces. This is also the case for very large water soluble organic molecules. Hydrophobicity is another important factor governing adsorption to soil surfaces. Hydrophobic (lipophilic) contaminants tend to adsorb strongly to the soil's organic matter. Molecular charge and hydrogen bonding are given as the third and fourth major factors in adsorption. Organic molecules possessing an intrinsic positive or negative charge are adsorbed to cation or anion exchange sites on the soil particles. The hydrogen bond is where a hydrogen atom acts as a bridge between two electronegative atoms, one on the contaminant and the other on the adsorptive surface. The strength of adsorption of a contaminant is inversely related to its availability and mobility in the soil.

The majority of cases of soil contamination with these contaminants are the result of industrial land use, such as petrol stations, fuel and oil depots, et cetera, where tanks were leaking or spills occurred. However, river deltas are the natural environment where buried organic materials often are transformed into petroleum hydrocarbons so that low concentrations of these contaminants and their derivatives can occur naturally in the soil. Marine hydrocarbon seeps that are the main source of floating tar and tar in sediments are mentioned by Dragun (2007) and are not uncommon.

Organic lipophilic contaminants that enter the soil in much smaller quantities, but nevertheless are of great importance, include a wide range of pharmaceutical products, some of which are used in bulk in animal husbandry and are applied to the soil in manure (Kümmerer 2001, Kümmerer 2004).

2.3.1.3 Metals and Metalloids – Trace Elements

Metals and metalloids that are used as catalysts and pigments and for electrolytic plating and in other objects that are composed of metals are mostly in ionic form, but can also be in elemental form, as contaminants in the soil. The most common ones are: silver (Ag), arsenic (As), cadmium (Cd), chromium (Cr), copper (Cu), nickel (Ni), lead (Pb), tin (Sn), (vanadium) V and zinc (Zn). Their mobility and bioavailability depend on their speciation which in turn depends on the soil pH and on the concentration in the pore water of anions like chloride. Chloride ions are capable of complexing some metals to form highly soluble negatively charged anions, for example $PbCl_6^{2-}$. Thus, the presence of sodium chloride in pore water affects the mobility of lead. Copper can be complexed to dissolved soil organic carbon (DOC) and is then as mobile as the DOC. The solubility of the DOC, in turn, is affected by the calcium concentration in the pore water (Römkens 1994).

However, several metals and metalloids also occur naturally in all geological materials, rocks as well as soils, to varying degree, depending on their chemical affinity to various minerals, varying presence in magmas, varying enrichment by hydrothermal processes and varying history of weathering. Man-related soil contamination with most of the metals mentioned above generally poses a much greater environmental risk than natural geochemical accumulation.

2.3.1.4 Salts and Bases of Metal Alkaloids and Boron

Apart from trace elements, consideration must be given to other inorganic contaminants, particularly bases of metal-alkaloids and nutrients.

Sodium chloride (NaCl) is widely used in industrial and domestic settings, as well as for de-icing agents on roads. High concentrations of highly soluble salts present deleterious conditions to plants, whether the result of human activities or the salts have accumulated naturally. There is an extensive literature on this subject starting with a classic publication by the US Salinity Laboratory (1954) and the no less classic theoretical understanding of soil colloidal systems, according to the Donnan and Gouy Double Layer Theories (Babcock 1963). Also a lot of modern literature on the effects of salt in soil is available, including a plethora of modern handbooks, such as by Tanji (1990), and various publications from authors published in the FAO Series of Irrigation and Drainage Papers.

The most common contaminant within this category is sodium chloride, but other chloride salts, sulphate salts, carbonates and hydroxides can also profoundly alter and damage the soil, as well as destroy biological life in the soil. Soil salinisation is the most well known form of contamination by salt. Much salt turns up in sewage and other wastewater flows and if these are subsequently applied to soil in irrigation water, significant soil deterioration and losses of productivity will take place, along with other ecological damage.

High salinity causes high osmotic tension in the pore water, making it more difficult for plants to extract water from the soil. Moreover, high chloride concentrations can be toxic in their own right. Salinity is normally measured as the electrical conductivity (EC) in μS cm^{-1} in a saturation extract or, in some countries, in a soil-water mixture, usually at a ratio of 1 part soil and 5 parts de-ionised water. Less soluble salts, such as gypsum ($CaSO_4.2H_2O$), have a maximum solubility of the order of 20–30 me L^{-1} in non-gypsiferous soils and do not increase the EC of the pore water to unacceptable levels.

Irrigation with naturally saline river water or groundwater, similarly without adequate leaching and replacement of sodium by added calcium (gypsum), will salinise and destroy the soil as a medium for plant growth (Tanji 1990; US Salinity Laboratory Staff 1954). Accordingly, irrigation can cause detrimental contaminant-like impacts. Application of high concentrations of potassium salts to the soil, often as carbonate or of potassium or sodium hydroxide as caustic cleaning agents, likewise can destroy the soil structure, permeability and its capacity to support plant growth.

High concentrations of sodium or potassium salts can produce a bad impact on soils if leaching takes place with low salinity water and without replacement of the monovalent cations by divalent cations, e.g., Ca^{2+} through an increase in shrink-swell potential, an increase in clay dispersion and a decrease in structure stability and hydraulic conductivity.

In general, elements usually considered as nutrients, such as boron, can transform into a contaminant with respect to concentration and solubility. Boron is often found in soils at elevated concentrations in Australian soils in low rainfall regions (McBride's strongly alkaline, confined environments). In these cases the boron occurs as a water-soluble form. Total element analyses carried out for contamination assessments may also find elevated boron, but if a hot water extraction yields a low boron result, the boron will be in the soil as the mineral tourmaline, which is one of the most stable, weathering-resistant minerals in the mineral kingdom.

Man-made salinity is widespread in most dry regions of the world where irrigation is carried out. In Australia it also is very common where removal of the original vegetation of woodland or native scrub and perennial herbaceous ground storey plants has led to decreased evapo-transpiration, increased rainwater accessions to the groundwater, rising groundwater tables and capillary transport of salt to the soil surface. This process is called "dry land salting" to distinguish it from irrigation salting.

2.3.1.5 Nitrogen and Phosphorus

Similar to some compounds mentioned above, even macronutrients such as nitrogen and phosphorus may behave like a contaminant.

Nitrogen as a biologically vital element will occur in a very large number of natural organic materials. In mineralised form as ammonia, ammonium, nitrite or nitrate it is highly soluble in water and mobile, but in water ammonium acts like a cation and is subject to the same exchange processes as other cations. When adsorbed as exchangeable cation it is not mobile, but easily exchanged with other cations. All mineral forms are subject to a wide range of biochemical processes that can reduce the oxidised forms to N_2 (denitrification), oxidise the reduced forms (nitrification), and convert them to new organic contaminants.

Phosphorus as another vital element in living systems has a somewhat simpler chemistry in the soil, as in mineralised form it occurs only as the phosphate (PO_4^{2-}) anion in soil. Like nitrogen, it also occurs in organic contaminants. The phosphate anion is strongly adsorbed in most soils by the coatings of ferruginous material around soil particles, and it can form low-solubility compounds with aluminium hydroxide, iron oxy-hydroxides and calcium. The most common phosphate soil minerals are fluorapatite ($Ca_5(PO_4)_3F$), hydroxylapatite ($Ca_5(PO_4)_3OH$), variscite ($AlPO_4.2H_2O$) and strengite ($Fe^{3+}PO_4.2H_2O$) and all have very low solubility. Therefore, phosphate is generally only slightly mobile in soils which contain significant amounts of clay, the above mentioned oxy-hydroxides or calcium carbonate. The maximum solubility of these phosphates is when the soil pH is around 6.5 (Wolt 1994).

The most discussed aspect of excessive concentrations of N and P in the soil concern the off-site impact. An example in the case of nitrogen is the risk to the groundwater quality and to streams and lakes from the inflow of high N seepage, and in the case of phosphorus the risk of high-P contaminated soil particles washing off the land with runoff to adjacent open water bodies. However, there are also other risks. Clay-sized soil particles become increasingly dispersive and mobile with increased sorbed P (Giszczak et al. 2006). Thus they can move more easily within the soil and over the soil surface with runoff. Likewise, Zhang et al. (2003) found that colloidal iron oxides in sandy soil became much more mobile in the pore water as the amount of sorbed P increased and that other trace elements followed the trend.

Smil (2000) quotes a 1977 survey by Frissel and Kolenbrander (1977) which showed that continuing applications of manures and fertilisers to agro-ecosystems created annual surpluses of P in the soil of 40 kg ha^{-1} in the Netherlands, 30 kg ha^{-1} in France, 25 kg ha^{-1} in Germany and 10 kg ha^{-1} in England.

Excessive phosphorus, if plant-available, that is more soluble, can create deficiencies of iron, manganese and zinc for plants by forming poorly soluble phosphates of these important trace nutrients.

The global nitrogen cycle has been grossly altered by industry and agriculture (Vitousek et al. 1997). Much increased nitrogen inputs into the soil system has caused serious soil acidification, loss of calcium and magnesium and increases of nitrogen in the groundwater and open water bodies.

Excessive plant-available nitrogen, for example as nitrate, tends to accelerate the oxidation of organic matter in soils and reduce the quality of the soil as well as producing more greenhouse carbon dioxide. In ecosystems where the nitrogen content of the soil is limiting plant growth, adding nitrogen can lead to more sequestration of carbon through increased photosynthesis. Nitrogen toxicity often leads to stunted growth and abnormal deep green foliage. Specific ammonium toxicity causes roots to become affected and brown with necrotic tips, while leaves can develop necrotic lesions and chlorosis (Wong 2005).

In Australia, large regions have been adversely affected by the use of superphosphate to promote the growth and productivity of clovers, especially subclover (*Trifolium subterraneum*) on soils with low cation exchange capacity and low calcium. The increased production of nitrogen by the clover, and its conversion to nitrate when organic matter decomposes, led to increased loss of calcium when the nitrate is leached out. The soils in these regions often became so acid that the clovers would cease to grow and pasture productivity dropped dramatically, while the cost of liming vast areas became a limitation. The process of soil acidification that is affecting millions of hectares in Australia became a prominent agricultural issue in the 1980s and 1990s (Coventry 1992; Coventry and Slattery 1991; Helyar 1991; Helyar and Porter 1989). In terms of many metals, increased soil acidity tends to increase their mobility.

Britto and Kronzucker (2002) discovered that many plants are sensitive to very high ammonium concentrations in the soil and that these plants are pumping out the NH_4^+ after having at first absorbed it at high rates. The removal of the NH_4^+ from the cells requires a great amount of energy and this energetic cost is the cause of a toxic shock.

2.3.1.6 Contaminants from Hospital Effluents and Sludges Discharged on Soil

A special but probably widespread type of soil contamination occurs when either treated hospital wastewater or sludge from hospital and municipal sewage treatments plants is discharged on the soil. The application can occur in rural and urban areas. Accordingly, it should be taken into account in relation to the urbanized areas as well.

Kümmerer (2001) reviewed the range of pharmaceuticals that are excreted by patients and unused medicines that are disposed of in wastewater in hospitals. To this should be added the personal care products like shampoos and soaps that likewise find their way into the wastewater and other segments of the environment (Klaschka et al. 2004), such as soils, lake or sea sediments, groundwater and surface water. This section concentrates on general rules about the behaviour of hospital contaminants once they have been dispersed in the soil based on the available research papers consulted.

Halling-Sørensen et al. (1998) reviewed the occurrence, fate and effects of pharmaceuticals in the environment, concentrating on the aquatic environment including the groundwater. The majority of their literature sources dealt with the biological impacts on living organisms in media where the effects were quickly discovered, such as fish farms, or observed by purposeful testing on Daphnia, algae, various microbial organisms and agricultural crops. Thiele-Bruhn (2003) and Hamscher et al. (2004, 2005) surveyed the fate of pharmaceutical antibiotics in soils.

The behavior of pharmaceutical drugs in the environment has been summarized by Velagati as follows (*verbatim)*:

- Most pharmaceutical drugs are resistant to partial or complete mineralization through biodegradation. Small chain peptides and other natural products may be mineralized and are, therefore, an exception. Aerobic-anaerobic cycles of biodegradation may enhance biotransformation and, in combination with photodegradation, may enhance the potential for depletion.
- Most pharmaceutical drugs are resistant to degradation through hydrolysis.
- In general, pharmaceutical drugs are mostly very poorly soluble in water and have high partition and adsorption coefficients. These properties contribute to strong adsorption to organic matter and the potential for bioaccumulation and bioconcentration. They are very strongly bound to sludge in the Publicly-Owned Wastewater Treatment Systems (POWTP), and to soil. Such properties also make them unavailable for microbial degradation (aerobic and anaerobic biodegradation) and also for hydrolysis and photo-degradation.
- When bioavailable, pharmaceutical drugs appear to biotransform and mineralize in soils more rapidly than in water because of the diversity of microorganisms present in soils (fungi, bacteria, and actinomycetes). Poor biotransformation observed under aerobic conditions in the POWTP is likely to be due to the lack of such diversity (very few fungal and actinomycetes populations) in microbial populations.

Taken from "Behavior of pharmaceutical drugs (human and animal health) in the environment", by Ranga Velagati, PhD, Director, Pharmaceutical Manufacturing Support Group, Analytical-Biochemistry Laboratories, Columbia, Missouri. Published in Drug Information Journal, Vol. 1, pp. 715–711, 1997.

Some antibiotics, e.g. the macrolides, are usually very poorly soluble in water but some are completely water-soluble or dissociate at typical soil pH values. Once adsorbed to the soil's solid phase, desorption often shows a hysteresis, meaning the adsorption curve and desorption curves are different, with the latter showing incomplete desorption.

There is a range of chemical behaviours described in Thiele-Bruhn (2003), Hamscher et al. (2004), Hamscher et al. (2005) and summarised by Thiele-Bruhn (2003). Some important examples of chemical behaviours are:

- The epimers and metabolites of various antibiotics can react in soil quite differently from the original substances.
- Expandable three-layer[4] clay minerals possessed much stronger adsorptive power to aminoglycosides, tetracyclines and tylosin than did the non-expandable illite and the two-layer mineral kaolinite. Such adsorption could involve inter-layer adsorption.
- Soil organic matter also strongly sorbs antibiotics, but it also depends on the amount and its composition.
- The rate of adsorption of most antibiotics in soils is fast, with a specific time frame mentioned of several hours for efrotomycin and sulfonamides.
- Antibiotic potency is generally reduced by adsorption and hence desorption reactivates it. Even so, sorption does not totally and necessarily eliminate antimicrobial activity.
- Mobility and transport in the soil of antibiotics depends obviously on their water solubility so that those that have low solubility are strongly retarded. To the extent that there are large pores and preferential flow paths, antibiotics can move rapidly.
- Degradation of antibiotics in soils can rarely take place through hydrolysis (an abiotic process) and oxidative decarboxylation and hydroxylation (enzymatic processes). It has been found that adding manure or sludge with high numbers of micro-organisms can speed up biodegradation of the antibiotics.
- The degradation of most xenobiotics occurs faster and more completely when the soil is aerobic than when it is not, and ciprofloxacin, for example, did not degrade at all under anaerobic conditions.

A number of authors reported concentrations of antibiotics found in soils, mostly in soils that were fertilised with treated sludge from sewerage systems (Table 2.3). However, it should not be forgotten that antibiotics will occur naturally in soils.

---

[4] Clay minerals have layered crystal structures comprising octahedral sheets of $Al(OH)_3$ and tetrahedral sheets of Si_2O_5. In some, the sheets can move apart to let other ions or molecules enter.

Table 2.3 Concentrations of antibiotics encountered in agricultural soils

Author(s)	Medium	Compounds and concentrations found
Thiele-Bruhn (2003)	Agricultural soil	Tetracyclines 450–900 μg kg^{-1} Macrolides 13–67 μg kg^{-1} Fluorquinolones 6–52 μg kg^{-1}
Thiele-Bruhn (2003)	Soil fertilised with manure	Tetracycline average 199 μg kg^{-1} Chlortetracycline 7 μg kg^{-1} Sulfadimidine 11 μg kg^{-1}
Hamscher, Pawelzick, Höper and Nau (2004, 2005)	Sandy soil fertilised with liquid manure	Tetracyclines averages > 150 up to 300 μg kg^{-1} Sulfamethazine max. 11 μg kg^{-1}
Beausse (2004)	Sludge-treated soils, topsoil 0–2.5 cm	Ciprofloxacin and norfloxacin residual 250−300 μg kg^{-1}, 5 months after sludge disposal respectively 450 and 350 μg kg^{-1}

2.3.2 Adsorptive Behaviour and Specific Surface Areas

Soils have a major influence on the mobility of contaminants by virtue of their cation and anion exchange capacity, specific adsorption (chemisorption) through the formation of co-valent bonds between an element and the mineral surface, and solid state diffusion by the penetration of an element into the pore spaces of a mineral's structure (Dragun 2007).

In this section soils and geological materials are classified in terms of adsorptive behaviour and specific surface areas.

The cation exchange capacity (CEC) of soils depends greatly on the content and mineralogy of the clay fraction. Generally, younger and less weathered clays have much greater specific CEC than older and more strongly weathered clays. Where leaching can occur, weathering often results in a loss of silica from the soil. Thus, the 2:1 layer type aluminosilicate minerals like smectite, vermiculite may be considered the less weathered clay minerals, with kaolinite being a highly weathered clay, and illite occupying an intermediate position. Youthful soils on volcanic ash deposits often have poorly crystallised clay minerals that have a high CEC. Gibbsite (γ-AL(OH)$_3$) represents a highly weathered aluminium clay mineral. Along with the loss of silica by weathering, other much less soluble soil minerals are formed and concentrated in the soil, such as various iron, manganese, titanium and aluminium oxides, hydroxides and oxyhydroxides (Table 2.4).

The anion exchange capacity (AEC) of mineral soils is dependent on the size of the area along edges of clay minerals, where permanent positive charges exist, and on the amount of oxides and hydroxides of aluminium, iron, and manganese, which can generate both cation and anion exchange due to the adsorption of protons and hydroxyl ions, and hence on the pH of the soil (McBride 1994; Sposito 1989). Soils displaying both cation and anion exchange, depending on their pH, are "variable-charge" soils. For example, adsorption of Cr(VI) by aluminium and

Table 2.4 Metal oxides, oxyhydroxides and hydroxides found commonly in soils

Name	Chemical formula	Name	Chemical formula
Anatase	TiO_2	Hematite	$\alpha\text{-}Fe_2O_3$
Birnessite	$Na_{0.7}Ca_{0.3}Mn_7O_{14}.2H_2O$	Ilmenite	$FeTiO_3$
Boehmite	$\gamma\text{-}AlOOH$	Lepidocrocite	$\gamma\text{-}FeOOH$
Ferrihydrite	$Fe_{10}O_{15}.9H_2O$	Lithiophorite	$(Al,Li)MnO_2(OH)_2$
Gibbsite	$\gamma\text{-}Al(OH)_3$	Maghemite	$\gamma\text{-}Fe_2O_3$
Goethite	$\alpha\text{-}FeOOH$	Magnetite	Fe_3O_4

Sposito (1989)

iron soil colloids becomes increasingly strong with increasing acidity (Loeppert et al. 2003).

Finally, the organic matter fraction of soils also contributes a major portion of the soil's CEC by virtue of the negatively charged endings of these organic molecules, resulting from the dissociation of hydroxyls from –COOH-groups, while non-polar organic contaminants can bind to the soil organic matter. Dissolved organic carbon can complex certain metals such as copper and make them mobile (Römkens 1994).

The degree of adsorption is a function of the specific surface area of the soil as well as of the nature of that surface area, for example the mineralogy of any coatings that differ from the particle itself, and the presence of micropores in that coating.

It has been known for a long time that heavy metals are strongly adsorbed by sesquioxides (McKenzie 1980) and by soil minerals with strong cation exchange or anion exchange capacity. Many metals like barium, cadmium, manganese, strontium, and zinc occur exclusively or generally as cations in the soil solution and hence are affected by the soil CEC, but others like arsenic, antimony, selenium and vanadium usually occur in the soil as negatively charged oxy-anions and are subject to the soil anion exchange capacity. The strength of adsorption is affected by the pH of the soil, with higher pH increasing the amount sorbed by hematite and goethite as well as birnessite. Trivedi and Axe (2007) relate the sorption of metals to the intraparticle diffusion in hydrous metal oxides. Diffusion is a time-dependent process and in experiments with nickel and zinc it appeared that slow intraparticle diffusion of the metal ions was the active process. After 4 hours of contact time the metal ions exhibited a similar local structure to a system aged for 110 days. There is, however, still very much to be understood of these processes that decrease the sorption and increase availability of metals in soils. Figures 2.7 and 2.8 represent adsorption of metals on iron and manganese oxides

Sorption of metals also takes place in nature over the long term. For example, ferruginous nodules formed in soils by pedogenic processes often have elevated concentrations of trace elements, such as arsenic, cobalt, copper, chromium, nickel, vanadium and zinc. Geologic weathering processes that cause an accumulation of iron oxide-hydroxide similarly are accompanied by enrichment in these metals. Some examples of this are provided in the next section, because they can be considered to be inherited, as opposed to anthropogenic.

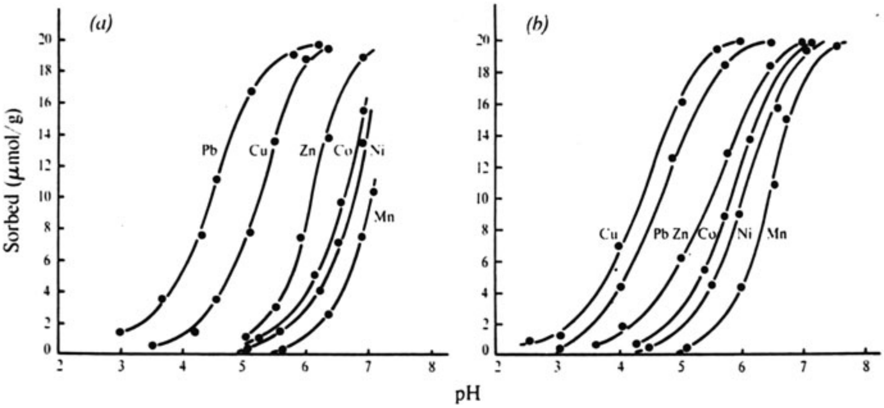


Fig. 2.7 Adsorption of heavy metals on iron oxides as a function of pH, 20 mmol/g added. (**a**) Hematite, (**b**) Goethite (McKenzie 1980)

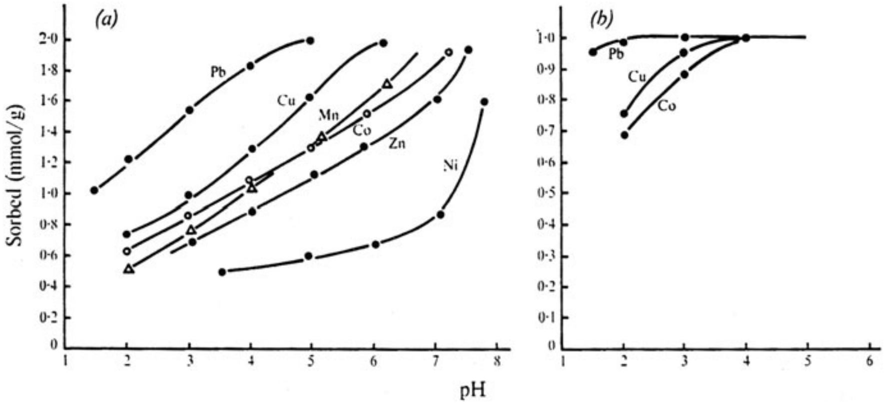


Fig. 2.8 Adsorption of heavy metals on birnessite as a function of pH, (**a**) 2 mmol g^{-1} added, (**b**) 1 mmol g^{-1} added (McKenzie 1980)

2.4 Contamination Potential

2.4.1 Soils of Deposited Material and Former Industrial Sites

In the presence of technogenic material, an enhanced contamination potential has to be taken into account (Nathanail and Bardos 2004). The level of contamination depends upon the kind of material, the amount of material deposited, and the degree of weathering. Table 2.5 presents the concentrations of some heavy metals being concentrated in different sorts of ashes. While the level in coal is low, ashes reveal higher values, in particular ashes derived from garbage incinerators. Moreover, the table also illustrates the differences in metal concentrations between the different kinds of ashes. It shows that metal concentrations in the silty fly ashes exceed clearly

Table 2.5 Cadmium, lead and zinc content (mg kg^{-1}) of coal and different sorts of ashes (from: different sources, e.g. Meuser 1996)

	Cadmium (Cd)	Lead (Pb)	Zinc (Zn)
Hard coal (raw material)	0.1	40	50
Bottom ash of coal power stations	2.1	82	72
Fly ash of coal power stations	14	1,080	990
Bottom ash of garbage incinerators	33	2,088	4,538
Fly ash of garbage incinerators	316	6,935	2.1%

the values of the coarse bottom ashes. Accordingly, it can be expected that different total concentration might lead to different mobile concentration between the technogenic materials.

In general, the texture influences the contamination level, since fine earth is able to adsorb more cations (e.g. cadmium, lead, zinc) than the coarse fractions. Therefore, some technogenic material consisting of silt or clay may be more contaminated. For instance, harbor sludge fields in Hamburg, Germany, show different heavy metal values, depending on the texture class (Fig. 2.9). On the other hand, it should be noted that areas consisting of mixtures between soil and coarse unweathered technogenic materials such as slag can show a higher contaminant concentration in the sandy and gravely fraction than in the silty and clayey.

In Table 2.6 the expected contaminant situation related to the different artificial components is summarized (Meuser and Blume 2004). The most important artificial component group is *construction rubble*, derived from buildings that have been removed. If decontamination of potentially problematic materials with contaminants and dismantling processes in a controlled manner is not possible, materials passing the demolition operations continuously contain contaminants, causing enhanced

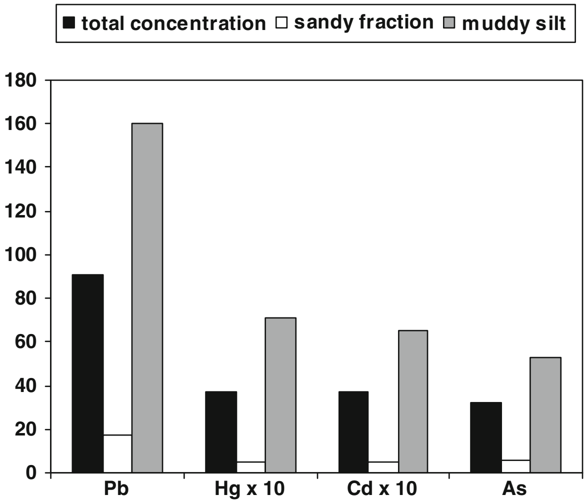


Fig. 2.9 Heavy metal concentrations (mg kg^{-1}) of harbor dredged sludges divided into different texture classes in Hamburg, Germany (unpublished)

Table 2.6 Contamination potential of technogenic components

Main component group	Contamination potential
Construction rubble	• mixtures generally higher contaminated than single components • problematical components: concrete based on fly ashes (Pb), asbestos containing debris (fibres), tubes containing debris (Cu, Pb), debris influenced by fire damages (PAH) • tar asphalt generally higher contaminated than bitumen asphalt (PAH, phenol)
Slag	• high heavy metal concentrations in metal works slag • high Cr and Ni concentrations in steel works slag • low heavy metal values in blast furnace slag, foundry sand, blast furnace slag, sand and pumice of iron works
Ashes	• fly ash generally higher contaminated than bottom ash • garbage incinerator ash generally higher contaminated than ash from coal power station • problematical components: dross, bottom ash (PAH)
Mining waste	• heavy metal contamination (ore mining waste) • radioactive emission (uranium mining waste) • strong acidity (hard coal mining waste) • salt leaching (salt mining waste)
Refuse	• problematical components: some plastics (heavy metals), wood (pesticides), organic garbage (methane formation) • generally indifferent contamination • occasionally high levels of industrial components

Meuser and Blume (2004)

migration by leaching and wind action (Nathanail and Bardos 2004). Besides, belowground structures like foundations are frequently not removed, for example during Brownfield redevelopment. Instead, former hollow structures have to be refilled or stabilized rapidly. Occasionally, even tanks, power and phone cables, and pipes are not removed (Genske 2003). Consequently, in some cases it may be problematical to distinguish between "soil" and foundation structures in situ. In areas under construction, often little natural soil remains, and the surface may be covered with debris from buildings that previously occupied the site. This situation may cause several difficulties, since, for instance, plaster (gypsum) causes concrete corrosion and disruption, and in dry conditions disturbed asbestos leads to a threat to human health, after inhalation (Bullock and Gregory 1991).

The contamination status is always changing, because deposited soils take part in *pedogenesis* in the same manner as natural soils. Results from pedogenesis may alter the mobility and availability of contaminants as well. Some typical features of pedogenesis that can be recognized in anthropogenically influenced soils are:

- humus accumulation leading to the formation of new A horizons;
- weathering processes, such as swelling-shrinkage, thawing-freezing, and acidification;
- reductomorphology caused by impeded water, human-made compactness, and reductive gases;
- carbonate and salt formation resulting in calcium-crusts and salt-crusts (in relation to some technogenic materials indicating high calcium carbonate concentrations as well as enhanced electrical conductivity);
- turbation by fauna, but also by swelling-shrinkage processes, e.g., in sludge fields.

However it should be kept in mind, that the soil weathering processes become less intensive because of higher temperatures (0.3–3.0°C higher), reduced freezing-thawing processes and other climatic effects in urbanized areas.

Apart from the deposited soils, *former industrial sites* can occupy originally natural soils. Nevertheless, their contamination potential may be very high, since there are specific contaminant losses during industrial processes in the course of time. A typical example is that of gas works areas, where one can find residues containing ferric ferro cyanides and coal tar enriched with benzene, toluene, xylene, phenol and PAH. Another example is scrap yards, where dismantling of vehicles and other machinery takes place, producing contaminants as PCBs derived from break-up of electrical equipment, and waste oil needed to run conveyor belts, sorting machines, etc (Bullock and Gregory 1991; Genske 2003). Besides, soil contamination of industry sites often results from handling, storage, and transport of potentially polluted products and by-products.

Additionally, in urban environment underground leakages of e.g. wastewater tubes and gas pipelines causing dispersion of soluble contaminants or reductive gases are another source of soil contamination. In case of leakage, different kinds of damage are possible. For instance, corrosion, swelling effects of loamy or clayey soils, traffic influence or subsidence present in mining areas are able to initiate losses of wastewater or gas. Thus, spills of liquids containing e.g. ammonia, sulphate, or heavy metals are typical results of leakages indicating contamination to soil. Furthermore, gases like methane derived from corroded gas pipelines may be vulnerable to explosion.

2.4.2 Additional Sources of Contamination

The introduction of technogenic materials is one mode that can contaminate soils, but additional contamination factors can overlap each other (Fig. 2.10). It is well-known, that natural soils without any visible disturbance or presence of anthropogenic artefacts can be highly contaminated due to different sources of contamination:

- atmospheric precipitation (dust deposition);
- flood occurrences in alluvial floodplains resulting in contaminated particle sedimentation during or after the flood;

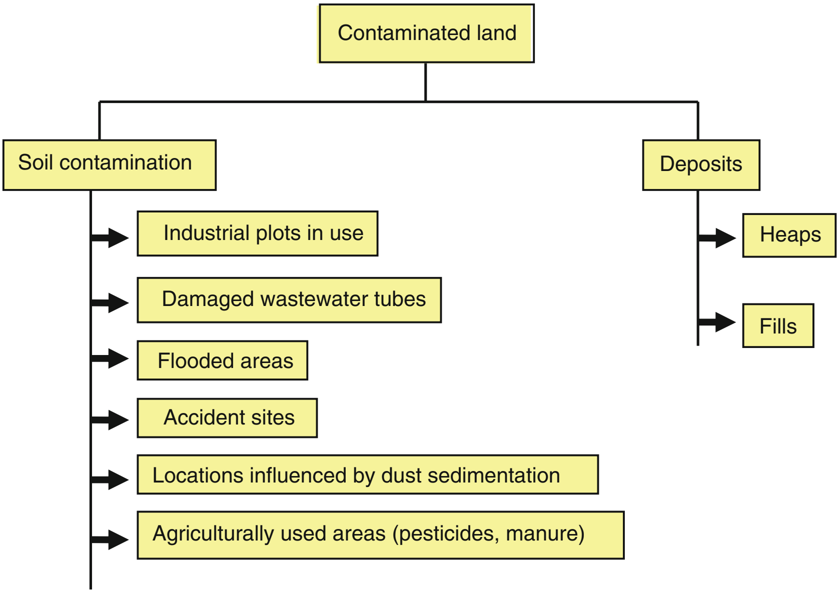


Fig. 2.10 Sources of contaminated sites

- accident sites with losses of problematical contaminants like mineral oils (leaking tanks and pipes of garages and petrol stations, accidental oil spills close to roads);
- agriculturally or horticulturally used sites, with possible application of pesticides as well as potentially contaminated manures like sewage sludge or composted municipal solid waste.

Most of the causes may lead to a depth gradient indicating decreasing contamination with increasing soil depth. Soil contamination often does not have visible features except for e.g. oil phases (accident sites) or artifacts (floodplains). Apart from the last example, the contamination source is located in the upper usually humic part of the profile connected with a high contaminant adsorption potential.

The *atmospherically determined soil contamination* involves both solid particles (particulate matter) and liquid droplets. In particular, atmospheric dust deposition plays an important role in districts with heavy traffic. On the one hand, with increasing distance to the road the decrease of deposition is rapid. On the other hand, in areas with intensive traffic, particularly the city center, an accumulation of dust on a large scale should be taken into account.

Dust deposition can be responsible for very high heavy metal concentration as examples in United Kingdom demonstrate (Table 2.7). The values in agglomerations like London are higher than in all studied locations of the United Kingdom. Very large concentrations, however, were found in mining villages, where dust development and enhanced geologically based background values have to be combined (Bullock and Gregory 1991).

Table 2.7 Lead and zinc concentrations in survey locations in United Kingdom (geometric means, mg kg^{-1})

Location	Depth (cm)	London districts	Mining villages*	All study locations in U.K.
Lead				
Garden soils	0–5	654	5,610	266
Vegetable plot soil	0–15	571	8,730	270
Public garden soil	0–5	294	3,030	185
Zinc				
Garden soils	0–5	424	9,340	278
Vegetable plot soil	0–15	522	8,750	321
Public garden soil	0–5	183	none collected	180

* Derbyshire (Pb), Shipham (Zn)
Bullock and Gregory (1991)

Accordingly, in different locations it has been found that the contaminants show a differentiation between urban and suburban areas. In particular towns in the USA, the top soils of urban agglomerations reveal higher contaminant concentrations than in the periphery of these locations (Fig. 2.11) (Craul 1992). Where the land surface is sealed, the contaminant concentrations are able to be lower than in fallow or vegetated areas, because dust was swept away or blown away by wind in dry conditions (Prokofieva and Gubankov 2000).

It should be taken into account that the material used for sealing purposes, in particular the load-bearing sub-base, is able to consist of problematical material causing subsoil contamination, if weathering processes are going to continue. In the two French cities La Teste and Le Mans, however, no soil contamination underneath the road was found in spite of their construction using municipal solid waste incinerator bottom ash, as the alkaline pH value had strongly reduced heavy metal solubility (Francois 2000).

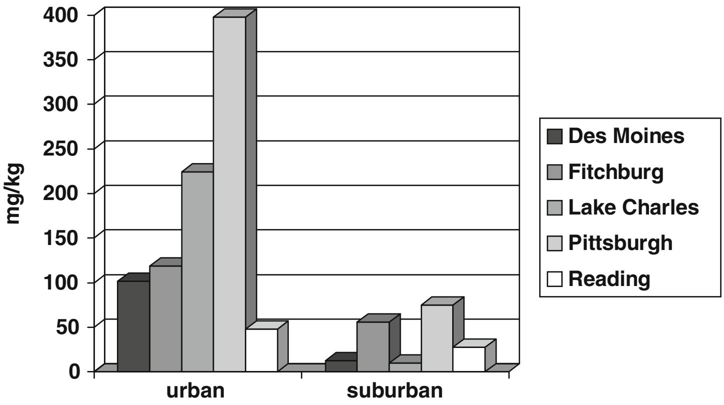


Fig. 2.11 Pb concentrations (mg kg^{-1}) in lawn topsoils of five U.S. towns depending on urban and suburban locations (Craul 1992)

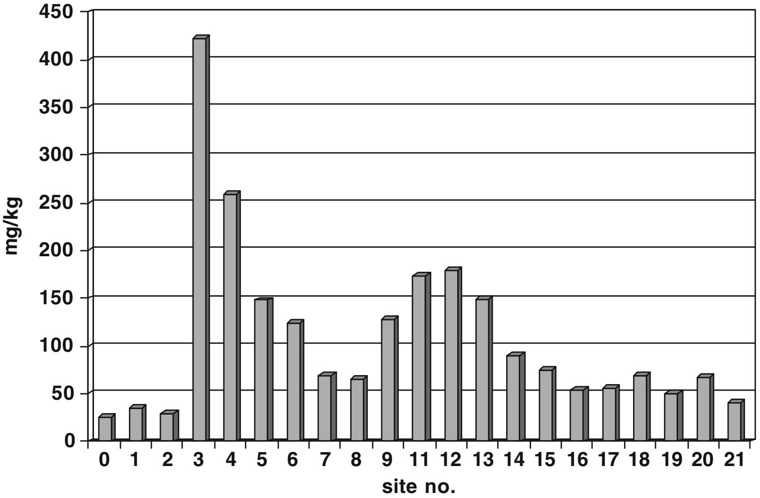


Fig. 2.12 Pb river sediment concentrations (mg kg^{-1}) in the heavily industrialized area of Pernik, south of Sofia, Bulgaria (Meuser et al 2008)

Except for atmospheric reasons, some soil contamination is able to be explained by geogenic causes in the line of background values. The size of the geologically caused soil contamination depends on the geological formation and ranges consequently from small to large areas.

In contrast, soil *contamination in alluvial floodplains* generally shows a linear distribution with distinct lengths and widths. As an example, the Strouma River in Bulgaria reveals a different contamination status, due to the possible sources of contamination (Fig. 2.12) (Meuser et al. 2008). Discharges located at a steel works (site No. 3 and 4 of Fig. 2.12) enhance lead values of the sediments abruptly, a second peak is visible entering the city centre (sites No. 11, 12, 13 of Fig. 2.12). Finally, the agricultural land and wetlands (behind No. 16 of Fig. 2.12) may reduce the lead values significantly.

Typical contamination patterns show a systematic relationship with the depth gradients of contamination. Contamination associated with dust sedimentation indicates a decreasing tendency with depth. Soils contaminated by parent material behave in the opposite direction. Soils influenced by flood and historically determined contamination generally show several contaminated layers in different depths.

2.5 Chemical Characteristics with Reference to Contaminated Sites

The presence of technogenic materials in soils may influence the chemical soil conditions considerably. Most of the calcium-enriched material is strongly alkaline, raising the *pH value* of the soil matrix. Construction rubble ranges between pH 6 and

8, slag and ashes between pH 7 and 11. Systematic topsoil investigations in three European cities, Ljubljana (Slovenia), Seville (Spain), and Torino (Italy), confirmed the neutral to alkaline reaction due to calcareous materials, since values higher than 7.0 have been found (Biasioli et al. 2008). Coal gangue, however, tends to acidification. Even pH values below 3.5 can be expected, which results in high heavy metal mobility (Meuser and Blume 2001).

The *carbon (C) content* of material such as ashes, tar asphalt, and coal remnants is very high, due to the inorganic carbon fraction that includes high aromaticity. Consequently an enhanced adsorption potential for organic contaminants should be taken into consideration (Abelmann et al. 2005). This inorganic carbon fraction does not consist of $CaCO_3$ only. Besides, inorganic carbon residues such as soot and fly ash determine the inorganic carbon fraction predominantly. Consequently, soils with a high percentage of inorganic carbon are easily detectable, because they have a wide C/N ratio. For instance, while the topsoil of an anthropogenically unaffected Chernozem showed C/N ratios of approximately 11, the topsoil of a neighbouring chernozem with strong lignite coal dust sedimentation close to a colliery had C/N ratios of 40 (Zikeli et al. 2002).

One of the most important filling materials is *household and commercial waste*. In particular, in countries where waste management is poorly developed, waste dumps are wide-spread. In India and Pakistan, for instance, the waste disposal can be qualified as problematical. The reasons for this are low collection capacity, in particular outside of big cities, the lack of landfill management, reduced recycle activities associated with street peddlers and rag-pickers, waste dispersion through animals (cattle), a low awareness in relation to waste problems, absence of garbage incineration (restricted to private companies, hotels, hospitals, only), and a low level of composting technology. It has been found that in urban areas, waste deposits apply often to the near-surface soil horizon, which contains high amounts of non-degradable waste, particularly plastic bags, hazardous industrial waste, and cattle waste (Panhwar 2000). Because of the high costs of mineral fertilizer in Ouagadougou, Burkina Faso, the inhabitants use waste to improve soil conditions (increasing soil depth from 15–20 to 30–40 cm, increasing humus content from 0.9 to 2.5%) for horticultural purposes (e.g., growing tomato, onion, salad, pepper). Coarse elements such as bottles, plastics and batteries are not sorted out, causing contamination of the degradable fractions (Thiombiano and Gnakambari 2000). Investigations applied to another African city, Dar es Salaam, Kenya mentioned the high plant uptake of contaminants by vegetables in suburban area and in old dumpsites (Luilo 2000).

Typical soils of deeper waste deposits show *reductomorphological features*, associated with the generation of reductive gases like methane and hydrogen sulphide. The reduction may influence the mobility of some heavy metals, causing accelerated leachate concentrations, possibly combined with groundwater contamination dependent on the hydrogeological conditions.

Chemically, in particular in the vicinity of the city centre, deposits consisting of organic residues (household waste) result in a relatively high *nutrient status*. The example of Nanjing, China, dealing with phosphorus, impressively reveals that with

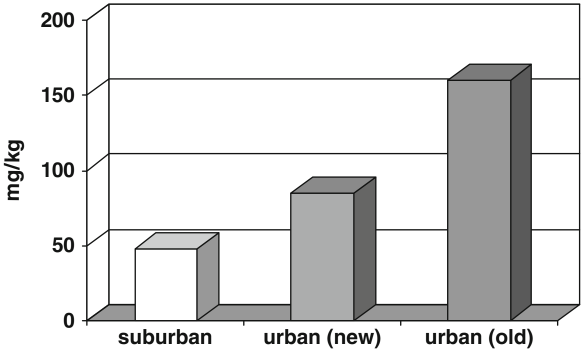


Fig. 2.13 Phosphorus concentrations in suburban and urban areas of Nanjing, China (Yuan et al. 2007)

increasing distance to the city centre the phopsphorus content decreased (Fig. 2.13). The suburban areas indicating a built-up period with current urban extension have lower soil P values in comparison with new (after 1947) and old (ancient times until 1947) urban terrains, consisting mainly of anthropogenic deposits (Yuan et al. 2007).

Another problematical technogenic material refers to the *coal mining waste*, the coal gangue, usually deposited in large heaps aboveground. There are several points of consideration in association with environmental damage. Firstly, dust resulting from the weathering processes near the surface can be blown away, affecting neighbouring sites. Second, slope instability can cause damage through land slides, in particular in heaps with steep slope gradients and a huge percentage of bare soil surfaces. The most important problems, however, are the oxidation processes after depositing sulfidic materials leading to oxidation of pyrite ($4FeS_2 + 15O_2 + 10H_2O \rightarrow 4FeOOH + 8H_2SO_4$). The resulting leachate of acids (acid mine drainage AMD), influencing the mobility of heavy metals which are able to contaminate groundwater and surface water, must be taken into consideration. Simultaneously, this exothermal chemical reaction may cause spontaneous ignition. In case of old coal mining heaps completely vegetated with shrubs and trees based on natural succession, the vegetation may catch fire aboveground.

The process of strong *acidification* is well-known in relation to the sludge lagoons, containing dredged harbour or stream sediment as well. As long as waterlogged conditions prevail, sulphur and iron remain in reduced form. But during drying, oxidation can take place producing sulphuric acid, and accelerating heavy metal mobility (Brady and Weil 2008). Thus, sites consisting of coal gangue or dredged harbour sludge tend to acidify. In industrialized areas this kind of acidification and the acid deposition from the atmosphere often decreases the soil pH value enormously.

A special kind of deposited material is sludge derived from raw material extraction like ore mine tailings. After separation processes (flotation, gravity separation, solution mining) it is necessary to dispose of acid waste water, using tailing ponds or lagoons. The tailing ponds may dry quickly. After drying, dust accumulated from heavy metals develops at the surface, is subjected to wind erosion, and sometimes contains contaminated agents like cyanides (Genske 2003). The *sludge fields* create bad conditions for plants due to contamination as well as physical impacts such as high soil density and platy soil structure. Fly ash fields consisting of 40% silt and 60% fine sand and indicating an alkaline pH value, underline the same negative impacts, deflation in dry conditions and strongly reduced root penetration resulting from the cement-like substance after reaction with lime (Bullock and Gregory 1991).

2.6 Physical Characteristics with Reference to Contaminated Sites

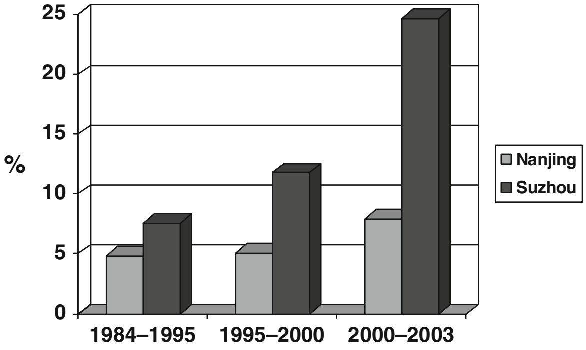
In urban areas, a lot of deposited soils are covered with sealing material to make an urban lifestyle possible. Therefore, apart from the areas covered with buildings, a high percentage of areas are covered with asphalt, concrete, or alternative material with a high load bearing capacity. In regard to town planning processes, in particular de-sealing and Brownfield redevelopment, some soils may be altered, e.g., excavated and transported off-site. Thus, the currently sealed surfaces can also be rehabilitated, and consequently should be taken into account in relation to contaminated site assessment.

The degree of *sealing* depends upon the utilization of the site. The degree of sealing in construction areas ranges between 35 and 80% and at roads and squares the values are higher than 60%. Results from investigations in the Southeast district of Moscow have revealed that industrial sites are usually >75% sealed, followed by residential areas (50–75%) and the natural areas (urban forest, agriculture) (<25%) (Stroganova et al. 1998).

Even outside the sealed areas *water infiltration* is often reduced, since the surrounding areas are influenced by compaction due to trucks and other vehicles, in particular caused during the construction phase. The compaction results from soil displacement and transportation and destroys soil structure. Moreover, the revitalization process of the soil is reduced, due to the low organic matter content as structure-forming agent. The reduced humus content results from mixing processes involving both topsoil and subsoil. The bulk density of the soil in refilled areas is increased by the passage of heavy traffic. In New York's Central Park, the bulk density in subsoil varies between 1.52 and 1.96 g cm^{-3} (Bullock and Gregory 1991). Still higher values are feasible, if the use of the sites requires drainage and high load-bearing capacity such as playgrounds and footpaths.

Craul (1992) mentioned water infiltration rates of 6.6 cm h^{-1} (sidewalk areas), 19.6 cm h^{-1} (baseball fields), 0.9 cm h^{-1} (areas with small slabs), while unsealed sites had values of 179.9 cm h^{-1} (abandoned pineapple field) and 61.7 cm h^{-1} (golf course). In contrast, deposits of technogenic materials like construction rubble or slag can have a very high saturated hydraulic conductivity, explained by their typical skeleton structure, leading to increased pore capacity between and within the material (e.g., pores in brick and slag) (Meuser and Blume 2004).

Fig. 2.14 Yearly increase rate of the urban area in Nanjing and Suzhou, China (Zhang et al. 2005, 2006)



A considerable portion of the deposited soils are sealed for building or road construction purposes. The expansion of city areas connected with *land consumption*, and consequently with losses of natural soils, is a well-known phenomenon. In the developing countries like China and India, these tendencies have reached spectacular dimensions. The development may continue in the next decades, since the degree of urbanization (people living in agglomerations with more than 150,000 inhabitants) has only achieved 43% in China, while in industrialized countries the values can be 92% (United Kingdom), 80% (USA), and 81% (Japan). Generally speaking, the development in the Chinese big cities such as Nanjing and Suzhou is faster than we have ever seen before (Fig. 2.14). The losses of soils do not include only less fertile soils with shallowness and high stoniness, fertile soils are involved as well. In Nanjing 14.5% of fertile agriculturally used soils (Argosols, Cambisols) were sealed, while the anthropogenically changed Anthrosols occupy 6.8% only (Zhang et al. 2005, 2006).

Apart from the completely sealed areas, deposited soils between buildings (usually used as lawns or small parks) show *physically damaged soil structure*, caused by soil handling. After deposited soils are placed, the normal organic matter cycle is interrupted (litter raked or swept away, only individual or groups of plants are cultivated, reduced biological activity) leading to a constrained vegetation cover. In turn, this tends to surface crusting (thin layer of washed-in particles), reduced water infiltration, and accelerated erosion (rill, gully) along anthropogenically overused lines like pedestrian foot paths, especially in parks, playgrounds, and sports fields (Craul 1992). The reason for enhanced wind erosion in areas where construction activity occurs is removal of vegetation, excavation and unprotected stockpiling of different materials. Deposits and removal of vegetation are combined with a high percentage of bare soils. Specific *indicator species* such as *Artemisia vulgaris* and *Solidago canadensis* living in such conditions can be found. Also specific indicator plants that have a high tolerance to heavy metals are of importance in relation to special deposits like ore deposits, where species tolerant to heavy metals (for example *Festuca ovina*, *Silene vulgaris*, *Agrostis tenuis*) can grow (Genske 2003).

Moreover, problematical physical impacts have to be considered in subsurface mining areas, some kinds of quarries, and pumping facilities of oil and gas.

The extraction processes leave voids behind that may cave in slowly leading to a long-term overburden subsidence. Subsidence can lead to collapsed buildings and wetted depressions creating new ponds or lakes. Additionally, subsurface mining, like hard coal mining, requires sites used for mining waste depositing. In the past, they were constructed as barren cones or flat-topped heaps. Nowadays, spoil heaps with natural-like topography and full vegetation are created. Analogously, surface mining, e.g., applied to lignite coal mining, means a greater extent of Greenfield consumption combined with changes of the morphology based on excavation and refilling operations. The terrain becomes completely devastated, creating horrible scars in the landscape (Genske 2003).

2.7 Case Studies

In most developed countries environmental assessments that include soil and groundwater analyses not infrequently turn up elevated concentrations of heavy metals and metalloids or other trace elements, like boron, that exceed the official Soil Quality Standard for defined ecological and human health impacts and hence create concerns about human health and ecological risk. In many cases these concerns are unwarranted because the contaminants are part of the natural geochemistry and inert. Risk Assessment would benefit greatly if high level soil science is called upon to analyse these situations. Unfortunately, there are few soil scientists involved in this work. A few examples may illustrate this.

2.7.1 The Soil as a Chromatogram – Barium

The expansion of Melbourne's urban land to the west required an environmental assessment of land that had never been used for anything else than grazing sheep and cattle. All of the western area is in McBride's "Strongly alkaline, confined environments" where basalt-derived soils are dominated by smectite clays and Na, Ca, and Mg salts accumulate. Most of these heavy clay soil profiles have slightly acidic top soils, but pH rapidly increases with depth. Calcium carbonate as soft accretions or hard nodules are ubiquitous at about 0.5–0.6 m depth.

The 2006 soil assessment (Wrigley et al. 2006) was carried out by grid sampling with more than 100 auger holes of a 15 hectare site near Deer Park, an outer suburb, with samples taken as per official prescription at the soil surface 0–0.1 m, 0.5 m depth and 1.0 m depth. A number of the 0.5 and 1.0 m samples showed elevated barium concentrations varying between 320 and 900 $mg \cdot kg_{dw}{}^{-1}$, thus exceeding the National Environmental Protection Measure (NEPM) for Ecological Impact Level (EIL) of 300 $mg \cdot kg_{dw}{}^{-1}$. The whole process of preparing the site for residential development came to a sudden halt, while in the meantime the cost of interest on the borrowed capital continued unabatedly.

Barium is likely to concentrate in intermediate and acid magmatic rocks, linked with alkali feldspar and biotite, although in this case the barium is derived from mafic (or "basic") rocks. Commonly, the concentration range is 400 $mg \cdot kg_{dw}{}^{-1}$ to

1,200 $\mathrm{mg{\cdot}kg_{dw}}^{-1}$ (Kabata Pendias 2001). World average natural abundance of barium in basaltic rocks is 250 $\mathrm{mg{\cdot}kg_{dw}}^{-1}$, but some basalts have much higher barium concentrations, e.g., tholeite basalt in north-east Ireland contains 1,350 $\mathrm{mg{\cdot}kg_{dw}}^{-1}$ of barium (Krauskopf 1967). As far as is known, there is no systematic survey of Victorian basalts that provide data on barium concentrations in these rocks and hence it is unknown what barium concentrations may be expected in the rocks from which these soils have developed.

In soils in the natural environment, barium will generally occur as barium sulphate ($BaSO_4$) due to the ubiquitous presence of sulphate ions. Some paint pigments contain barium salts and thus barium concentrations can be elevated on derelict land (Bridges 1987). Barium is also associated with radio-active fall-out and nuclear waste. Barium sulphate is widely used as a safe tracer in medical practice due to the insolubility of the compound.

Barium molecules released from weathering are not very mobile and readily precipitate as sulphates and carbonates as long as sulphates and carbonates are available in the soil profile or can be applied as amendments. It must also be remembered that, as other rock constituents are lost by leaching in the soil formation process, those constituents that are less likely to be lost become proportionally more abundant in the residue. The soil barium concentrations derived from laboratory tests, although exceeding the Soil Quality Standard (NEPC 1999) were within the typical range measured from soils in their natural state around the world. The elevated concentrations recorded at this site are associated with depths of between 0.5 and 1.0 m in well-defined soil profiles that possess a rock basement with little colonisation of plant roots beyond 450 mm due to the presence of a calcareous clay-dominated B-horizon. The barium accumulations also occur in the profile where free lime is omnipresent and the pH high, thereby limiting plant availability.

Barium precipitates in the presence of sulphate (${SO_4}^{2-}$) to form barite ($BaSO_4$), which is extremely insoluble ($Ksp = 1.3 \times 10^{-10}$). Thus sulphate can be employed to limit the actual level of risk to health and control mobility (Lehr et al. 2002), so that gypsum can be used as an amendment to limit barium availability.

Basic soil science would predict that the barium in the soil would be in an insoluble form, because all soils that have a living biosphere must contain sulphate from decomposed proteins and rainfall accessions. However, like lime, the barium concentration could be a function of depth with its main accumulation somewhere down the profile. Two pits were dug with an excavator for detailed sampling of the profile in layers of 10 cm thickness. The soil profile exposed by the first pit was between 1 m and 1.2 m deep, the excavation limited by weathered rock floaters of vesicular basalt. Layering in the profile was well-defined, with 600 mm of red-brown clayey-loam (A-horizon), overlying a calcareous grey heavy clay (B-horizon). The latter was strongly structured and showed “slickensides”, providing abundant evidence of shrinking and swelling with fluctuations in soil moisture content. The B-horizon was flecked with lime and underlain by an extensive calcareous layer about 100 mm thick. The profile exposed by the second pit had better internal drainage, with more pronounced layering and more duplex in character. Additionally, the rock was closer to the natural surface and the B-horizon thinner.

Chemical analysis of the samples for pH, EC, water soluble sulphate and concentrated HCl extractable barium showed that all of these parameters possessed a depth function (Fig. 2.15). That pH, EC, and percent exchangeable Na display a depth function in these clay soils was known a long time. A zone of calcium

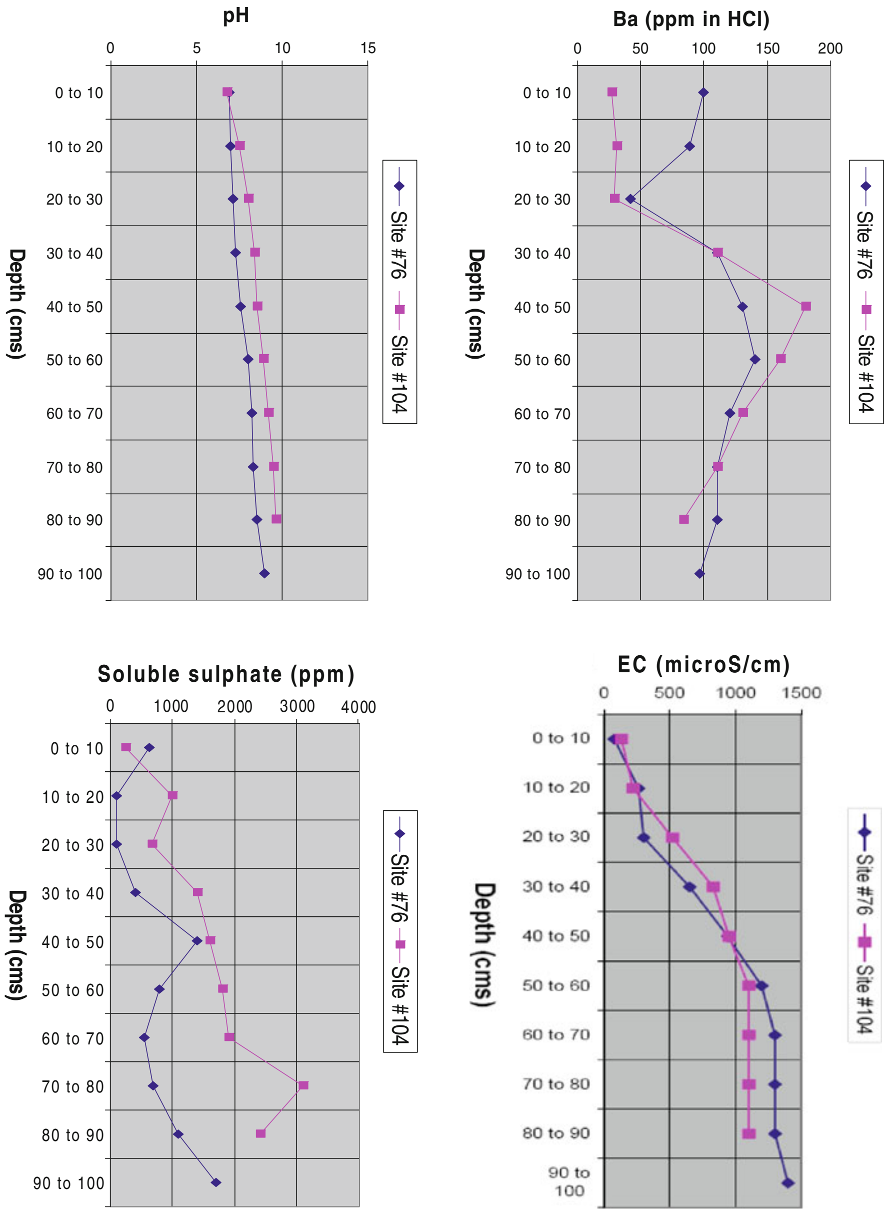


Fig. 2.15 Depth functions of some soil chemical parameters in basalt-derived heavy clay soils (Wrigley et al. 2006)

carbonate accumulation is often found below 0.5–0.6 m depth. Hence it was hypothesised that barium might behave in the same way in the form of a low-solubility contaminant.

Chemical analysis indicated that the molar ratio of water soluble SO_4^{2-} to extracted barium was of the order in the range of $1–10^{-5}$ to $1–10^{-6}$ and therefore the barium would all be in inert form as barite ($BaSO_4$). Notwithstanding its low solubility, over a very long geological period the barite had distributed itself over the soil profile in the form of a natural chromatogram. It was concluded that the barium was far too insoluble to present a risk to human health and the ecology and as a result the development was able to go ahead.

2.7.2 *Arsenic in Weathered Rock at the New Victorian Museum, Melbourne*

During excavations for underground car parking and an underground cinema for the new Museum Victoria a number of samples returned slightly elevated total arsenic concentrations in excess of the 20 $mg \cdot kg_{dw}{}^{-1}$ Soil Quality Standard for "uncontaminated" soil. The deeply weathered folded and faulted strata of shale, siltstone and sandstone of Silurian age were exposed in the excavation and showed strongly kaolinised and ferruginised zones. These were sampled separately and subjected to analysis of total iron and arsenic, citrate-dithionite extractable iron and arsenic (Holmgren 1967), followed by determination of total arsenic in the residue. Thus it could be proved that the bulk of the arsenic was generally liberated along with the extractable iron, and that the kaolinised zone possessed very little arsenic, although the high iron zone did not always have high arsenic. An extraction with a buffered solution of acetic acid and sodium acetate at pH 6 according to the Toxicity Characteristic Leaching Procedure[5] (TCLP), which is a routine extraction procedure in Victoria to model availability to plants, could not extract any arsenic from the samples.

Due to the sometimes extremely high iron contents in the rusted strata, the standard citrate-dithionite extraction was not able to liberate the bulk of the iron and therefore, presumably, the bulk of the arsenic. These analyses were part of a commercial environmental assessment and therefore could not be pursued with greater scientific thoroughness.

Table 2.8 summarises the laboratory data. Note that the TCLP extraction employs a mild extractant, a buffered sodium acetate – acetic acid solution at pH 6. Once again, the evidence provided by a soil chemistry-based assessment lifted the concern

---

[5] The TCLP, or Toxicity Characteristic Leaching (not Leachate) Procedure is designed to determine the mobility of both organic and inorganic analytes present in liquid, solid, and multiphasic wastes. The TCLP analysis simulates landfill conditions. Over time, water and other liquids percolate through landfills. The percolating liquid often reacts with the solid waste in the landfill, and may pose public and environmental health risks because of the contaminants it absorbs. The TCLP analysis determines which of the contaminants identified by the United States Environmental Protection Agency (EPA) are present in the leachate and their concentrations.

Table 2.8 Correspondence of As to iron accumulation, but not all iron-rich strata have much As

Depth (m) and sample #	Total As (mg kg^{-1})	Total Fe (mg kg^{-1})	TCLP As (mg L^{-1})	Total As after C/D (mg kg^{-1})	As in C/D extr. (mg kg^{-1}) (%)	Fe in C/D extr. (mg kg^{-1}) (%)
9.8 (C)	4.0	20,000	<0.001	1.8	2.2 (55)	– (28)
10.2 (C)	4.5	178,000	<0.001	3.6	0.9 (20)	– (7.8)
10.5 (C)	5.0	24,000	<0.001	2.8	2.2 (44)	– (24)
11.0 (C)	7.5	180,000	<0.001	4.5	3.0 (40)	– (7.8)
2.0 (V)	13	20,000	<0.001	2.3	10.7 (82)	– (41)
2.5 (V)	40.0	36,000	<0.001	24.0	16.0 (40)	– (47)
4.9 (V)	<2	2,000	<0.001	<2	<0.1	– (<0.1)
5.1 (V)	13.0	42,000	<0.001	6.1	6.9	– (36)
7.0 (V)	7.0	124,000	<0.001	1.3	5.7 (81)	– (25)
9.9 (V)	<2	680	<0.001	<2	<0.1	– (<0.1)
10.0 (V)	<2	6,800	<0.001	<2	<0.1	– (74)
10.1 (V)	5.2	124,000	<0.001	0.7	4.5 (86)	– (81)
10.2 (V)	11.0	114,000	<0.001	0.3	10.7 (97)	– (23)
3.0 (BH85)	25.0	38,000	<0.001	15.0	10.0 (40)	– (14)
3.2 (BH85)	8.5	55,000	<0.001	3.5	5.0 (59)	– (11)
3.6 (BH85)	2.8	5,100	<0.001	0.3	2.5 (90)	– (7.5)
6.0 (BH85)	4.0	25,000	<0.001	1.2	2.8 (70)	– (30)

Notes: All kaolinised strata have very little Fe (*light grey*) while all *dark grey* strata have high Fe; Lowest As samples are in kaolinised strata; *C* Coffey Partners core samples; *V* van de Graaff pit face samples; *BH85* CMPS&F core samples

and construction activities were allowed to proceed, with the excavated material being able to be disposed of cheaply as "Clean Fill".

2.7.3 Chromium in Soils

A large military base in Victoria was no longer needed by the Defense Department and the land was to be redeveloped for future residential uses. The area was largely situated on an extensive colluvial slope at the base of steep hills with gneiss and biotite gneiss bedrock. The area has a dry climate with a mean rainfall of some 740 mm per year and pan evaporation of around 1,600 mm yr^{-1}. In the winter and spring months, rainfall exceeds evaporation and some leaching occurs. The typical soils are Duplex soils with acidic redbrown B horizons (probably equivalent to Dystric Cambisols). An environmental audit of the soils by consulting environmental engineers found in a number of soil samples that chromium exceeded the EIL for that metal, so that in principle development could not go ahead. A follow up investigation was carried out by a consulting soil scientists (van de Graaff, Private Communication).

Table 2.9 Ionic radii of several metals and metalloids that frequently occur with iron and some that don't

Fe^{3+}	**Fe^{2+}**	K^+	**Cr^{3+}**	**Ni^{2+}**	**Co^{2+}**	**V^{3+}**	**Ti^{2+}**	Cu^+	**Cu^{2+}**	**Zn^{2+}**	Mn^{2+}	Pb^{2+}	Ba^{2+}	Sb^{3+}	Sn^{4+}
0.64	**0.74**	1.33	**0.63**	**0.69**	**0.72**	**0.74**	**0.76**	0.96	**0.72**	**0.74**	0.80	1.20	1.34	0.76	0.71

The National Environment Protection Council (NEPC) Environmental Investigation Level (EIL) for chromium is set at 1 $mg{\cdot}kg_{dw}{}^{-1}$ under the assumption the chromium will be hexavalent. However, other ANZECC[6] criteria are respectively 50 $mg{\cdot}kg_{dw}{}^{-1}$ for level B (*Environmental Soil Quality threshold standard necessitating further investigations*) and 400 $mg{\cdot}kg_{dw}{}^{-1}$ for NEPC for level B (*Soil Investigation Level for residential land with substantial vegetable garden contributing at least 10% of vegetable and fruit intake*). Cr^{6+} is not stable in soils which contain easily oxidisable organic matter, as is always the case in natural soils. Therefore the NEPC (EIL) is irrelevant. However, Table 2.9 indicates the Cr^{3+} would fit easily in any tri-valent iron compound in the soil and therefore is likely to be part of the crystalline structure of the ferruginous matter and equally insoluble.

The samples that proved to have elevated chromium concentrations invariably were the same where illuviation had caused the accumulation of iron along with clay in the B horizon. Or, in another case near Broadford, Victoria, where ferruginous nodules had developed at the interface between A and B horizons where seasonal waterlogging and drying occurs. Apparently the enrichment of the subsoil with chromium is merely a pedogenetical phenomenon. Biotite is also an iron-rich mineral and so it is likely the chromium is contributed by the biotite by isomorphous substitution for iron. Figure 2.16 illustrates that samples number 1–63, which are Duplex soil profiles with dark reddish brown B horizons, always have more chromium in the subsoil than in the topsoil. Samples number 64 and higher all relate to shallow skeletal soils (Leptosols), without subsoils overlying biotite gneiss bedrock.

Another proposed development site near Seymour also caused environmental doubts to arise when a number of soil samples showed high chromium. All these samples were from shallow topsoil which contained large numbers of ferruginous concretions (Fig. 2.17).

Although iron is not considered a contaminant, proper insight into soil chemical behaviour of chromium resulted in the decision to additionally measure the iron concentrations in soil, with the purpose of understanding the distribution and mobility of chromium in the soil. This approach proves that chromium in these soils at Bandiana and Broadford is practically inert and risk for the ecosystem or to human health is negligible.

---

[6] Australia and New Zealand Environment Conservation Council.

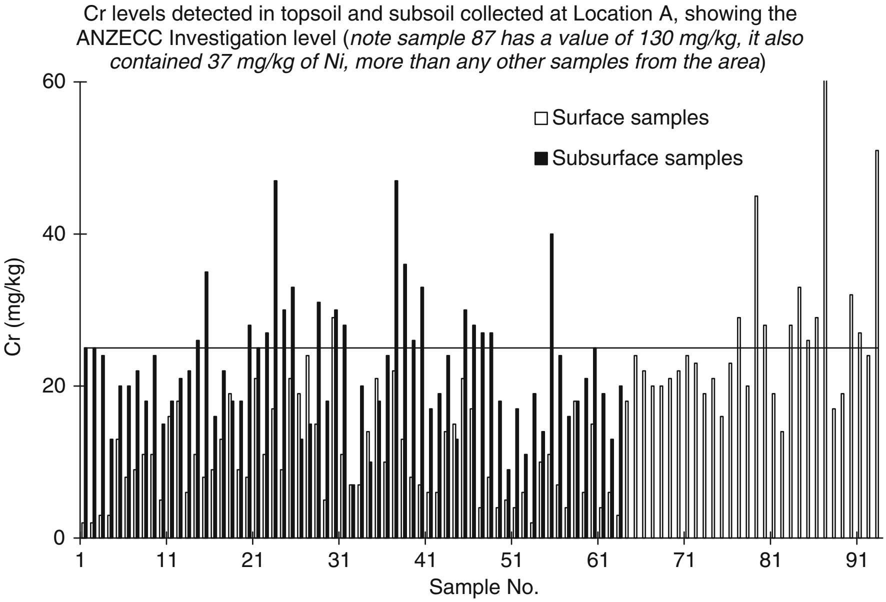


Fig. 2.16 Cr levels detected in topsoil and subsoil collected at the above locations, showing the ANZECC Investigation level for 2-composite samples (25 mg kg^{-1}). Note: sample 87 has a value of 130 mg kg^{-1}. It also contained 37 mg kg^{-1} nickel, more than any other sample from the area

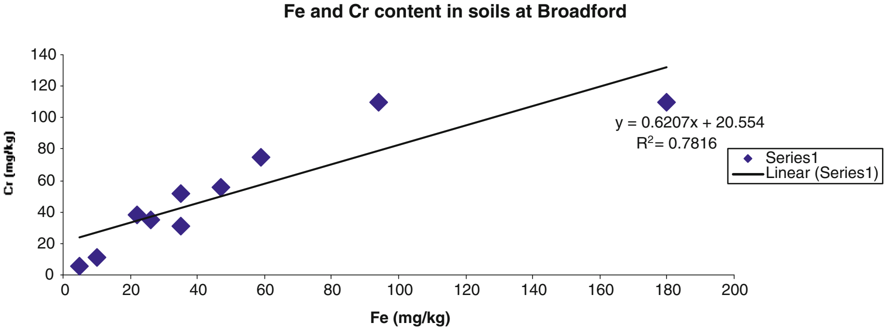


Fig. 2.17 Cr and Fe are closely related in soil samples containing many ferruginous nodules as shown by acid extractable Cr versus Fe contents collected at a location near Seymour, Victoria

2.7.4 Vanadium in Soils

Vanadium has accumulated preferentially in secondary ferruginous deposits in a weathering profile in sedimentary rocks in Heidelberg, a suburb of Melbourne to a level where some samples of this material exceeded the 50 mg·kg_{dw}^{-1} EIL for soils. Once again the redevelopment of the site came to a halt until the risk associated with the vanadium could be assessed.

On the basis of the ionic radii in Table 2.9 one can see that vanadium fits very well in crystalline oxidised compounds of divalent iron. Forms of iron oxy-hydroxides often combine a divalent and a trivalent form of iron, such as in magnetite, Fe_3O_4. According to the review of vanadium in soils by Jones et al. (1990), oxy-anions of V^{5+} are soluble over a wide range of pH and hence are the mobile forms in soils, but the degree of their mobility depends on physical and geochemical factors. Precipitation of less soluble forms of V may be caused by:

- the presence of reducing agents;
- local concentrations of elements forming insoluble vanadates (e.g. Ca);
- precipitation in the form of uranyl ($(UO_2)^{2+}$) cations;
- the presence of Al^{3+} or Fe^{3+} ions.

Of these factors the last one is richly present in the weathered seams of the local bedrock.

Chemical analysis of samples having visible iron oxides and having none, showed the vanadium is restricted to the ferruginous accumulations in these soils and the latter clearly have extremely low solubility. It could be concluded there is no soil contamination at this site and there is no reason for limiting the use of the land.

References

Abelmann K, Kleineidam S, Knicker H, Grathwohl P, Kögel-Knabner I (2005) Sorption of HOC in soils with carbonaceous contamination: influence of organic-matter composition. Plant Nutr Soil Sci 168:293–306

Babcock KL (1963) Theory of the chemical properties of soil colloidal systems at equilibrium. Hilgardia 34(11):417–542

Beausse J (2004) Selected drugs in solid matrices: a review of environmental determination, occurrence and properties of principal substances. Trends Anal Chem 23(10–11):753–761

Biasioli M, Grcman H, Kralj T, Madrid F, Diaz-Barrientos E, Ajmone-Marsan F (2008) Potentially toxic elements contamination in urban soils: a comparison of three European cities. Environ Qual 36:70–79

Brady NC, Weil RR (2008) The nature and properties of soils. Pearson Education International, NJ

Bridges EM (1987) Surveying derelict land. Clarendon, Oxford

Britto DT, Kronzucker HJ (2002) NH_4^+-toxicity in higher plants: a critical review. J Plant Physiol 159:567–584

Bullock P, Gregory PJ (1991) Soils in the urban environment. Blackwell Scientific, London

Coventry DR (1992) Acidifcation problems of duplex soils used for crop-pasture rotations. Aust J Exp Agric 32:901–914

Coventry DR, Slattery WJ (1991) Acidifcation of soil associated with lupins in a crop rotation in north-eastern Victoria. Aust J Agric Res 42:391–397

Craul PJ (1992) Urban soil in landscape design. Wiley, New York

Däumling T (2000) Results of the site survey of the public green of the city of Hamburg, Germany (in German). Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

Dragun J (2007) The soil chemistry of hazardous materials, 2nd edn. Amherst Scientific, Amherst, MA, USA

FAO (2006) World soil resources reports no. 103. International Union Soil Scientists Working Group WRB. FAO, Viale delle Terme di Caracalla, 00100 Rome, Italy

FAO-Unesco (1998) World reference base for soil resources. Information Division, Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla, 00100 Rome, Italy

Francois D (2000) Inspection of urban roads made with municipal solid waste incinerator bottom ash. Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

Fraters B, Boumans LJM, Prins HP (2001) Background concentrations of 17 trace elements in groundwater of the Netherlands (in Dutch). RIVM Report 711701017/2001. RIVM, Bilthoven, the Netherlands

Frissel MJ, Kolenbrander GJ (1977) The nutrient balances. Agroecosystems 4:277–292

Genske DD (2003) Urban land – degradation, investigation, remediation. Springer, Berlin

Giszczak M, van Leeuwen HP, van der Zee SEATM (2006) An effect of excessive phosphorus application on some properties of soil colloids corresponding with acceleration of their mobility in soil. Geophys Res Abstr 8:06826

Halling-Sørenson B, Nors Nielsen N, Lanzky F, Ingerslev F, Holten Lützhøft HC, Jørgensen SE (1998) Occurrence, fate and effects of pharmaceutical substances in the environment – a review. Chemosphere 36(2):357–393

Hamscher G, Pawelzick HT, Höper H, Nau H (2004) Antibiotics in soil: routes of entry, environmental concentrations, fate and possible effects. Ch. 11. In: Kümmerer K (ed) Pharmaceuticals in the environment. Springer, Berlin-Heidelberg

Hamscher G, Pawelzick HT, Höper H, Nau H (2005) Different behaviour of tetracyclines and sulfonamides in sandy soil after repeated fertilization with liquid manure. Environ Toxicol Chem 24(4):861–868

Helyar KR (1991) The management of acid soils. Develop Plant Soil Sci 45:365–382

Helyar KR, Porter WM (1989) Soil acidification, its measurement and the processes involved. In: Robson AD (ed) Soil acidity and plant growth. Academic, Sydney, pp 61–101

Holmgren GGS (1967) A rapid citrate-dithionite extractable iron procedure. Soil Sci Soc Am Proc 31:210–211

Jones KC, Lepp NW, Obbard JP (1990) Other metals and metalloids. Ch. 14. In: Alloway BJ (ed) Heavy metals in soils, Blackie, Glasgow & London; Halsted Press (John Wiley & Sons, Inc.), New York, 339 pp

Kabata Pendias A (2001) Trace elements in soils and plants, 3rd edn. CRC, Boca Raton, FL, USA

Klaschka U, Liebig M, Moltmann JE, Knacker T (2004) Potential environmental risks by cleaning hair and skin. Ch. 30. In: Kümmerer K (ed.) Pharmaceuticals in the environment – sources, fate, effects and risks, 2nd edn. Springer, Berlin Heidelberg

Krauskopf K (1967) Introduction to geochemistry. McGraw-Hill, New York

Kümmerer K (2001) Drugs in the environment: emission of drugs, diagnostic aids and disinfectants into wastewater by hospitals in relation to other sources – a review. Chemosphere 45: 957–969

Kümmerer K (ed) (2004) Pharmaceuticals in the environment – sources, fate, effects and risks, 2nd edn. Springer, Berlin Heidelberg

Lehr J, Hyman M, Gass TE, Seevers WJ (2002) Handbook of complex environmental remediation problems. McGraw-Hill, New York

Loeppert RH, Howe JA, Shahandeh H, Wei LC, Hossner LR (2003) Absorption and translocation of chromium by plants: plant physiological and soil factors. In: Naidu R, Gupta VVSR, Rogers S, Kookana RS, Bolan NS, Adriano DC (eds) Bioavailability, toxicity and risk relationships in ecosystems. Science, Enfield (NH), USA, Plymouth, UK

Luilo GB (2000) Vegetable gardening in urban Dar es Salaam and its associated risks: an overview on heavy metal contamination. Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

McBride MB (1994) Environmental chemistry of soils. Oxford University, Oxford

McKenzie RM (1980) The adsorption of lead and other heavy metals on oxides of manganese and iron. Austr J Soil Res 18:61–73

Meuser H (1996) Technogene substrate als ausgangsgestein der böden urban-industrieller verdichtungsräume (in German language). Schriftenr. Inst. f. Pflanzenern. Bodenkd., Kiel, p 35

Meuser H (2007) Wichtige Aspekte für Stadtböden bei der Bewertung der Bodenfunktionen gemäß §2 BBodSchG (in German language). Mitt Dt Bodenkdl Ges 110:513–514

Meuser H (2010) Contaminated urban soils. Springer, Dordrecht

Meuser H, Anlauf R (2007) Soil contamination of agricultural and horticultural sites in peri-urban areas. Paper presented at International Conference Sustainable Agriculture for Food, Bio-energy and Livelihood Security, Jabalpur, India

Meuser H, Blume H-P (2001) Characteristics and classification of anthropogenic soils in the Osnabrueck area, Germany. Plant Nutr Soil Sci 164:351–358

Meuser H, Blume H-P (2004) Anthropogene böden. In: Blume H-P (ed) Handbuch des Bodenschutzes (in German language). Ecomed, Landsberg, pp 573–592

Meuser H, Härtling A, Döpke G (2008) Pollutant dispersion of alluvial soils in the industrialized region of Pernik, Bulgaria. Paper presented at Eurosoil Congress 2008, Vienna

Mitchell RL (1965) Trace elements in soils. In: Bear FE (ed) Chemistry of the Soil. ACS monograph no. 160, Reinhold Publishing Corporation, New York

National Environment Protection Council (1999) National Environment Protection (Assessment of Site Contamination) Measure

Nathanail CP, Bardos RP (2004) Reclamation of contaminated land. Wiley, Chichester

Northcote KH (1960) A factual key for the recognition of Australian soils. Rellim Technical, Adelaide, South Australia

Panhwar F (2000) Soils observed in urban sites. Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

Prokofieva T, Gubankov A (2000). Sealed soils in urban soil cover. Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

Reimann C, de Caritat, P (1998) Chemical elements in the environment – factsheets for the geochemist and environmental scientist. Springer, Berlin Heidelberg New York, 398 pp

Römkens PF (1994) Interaction between calcium and dissolved organic carbon: implications for metal mobilisation. In: Abstracts 3rd international symposium on environmental geochemistry. Faculty of Geology, Geophysics and Environmental Protection, University of Kraków, Kraków, Poland

Smil V (2000) Phosphorus in the environment – natural flows and human interference. Annu Rev Energy Environ 25:53–88

Sposito G (1989) The chemistry of soils. Oxford University, Oxford, New York

Stasch D, Holland, K, Beck O, Stahr K (2000) Evaluation of soil resources in urban areas. Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

Stroganova M, Myagkova A, Prokofieva T, Skvortsova I (1998) Soils of Moscow and urban environment (ed). University of Essen and Lomonosov Moscow State University

Tanji KK (1990) Agricultural salinity assessment and management. American Society Civil Engineers, New York

Thiele-Bruhn, S. (2003) Pharmaceutical antibiotic compounds in soils – a review. J Plant Nutr Soil Sci 166:145–167

Thiombiano, L, Gnakambari X (2000) The amendment of Leptosols with urban wastes for crops and vegetables production in the area of Ouagadougou (Burkina Faso). Paper presented at 1st International Conference on Soils of Urban, Industrial, Traffic and Mining Areas, Essen, Germany

Trivedi, P, Axe, L (2007) Long-term fate of metal contaminants in soils and sediments: role of intraparticle diffusion in hydrous metal oxides. In: Hamon R, McLaughlin M, Lombi, E (eds)

Natural attenuation of trace element availability in soils. SETAC, Pensacola, FL; Brussels, Belgium, CRC Taylor & Francis, Boca Raton

U.S. Salinity Laboratory Staff (1954) Diagnosis and improvement of saline and alkali soils. Agricultural Handbook No. 60, U.S. Department of Agriculture

Vitousek PM, Aber JD, Howarth RW, Likens GE, Matson PA, Schindler DW, Schlesinger WH, Tilman DG (1997) Human alteration of the global nitrogen cycle: sources and consequences. Ecol Appl 7(3):737–750

Wolt JD (1994) Soil solution chemistry – applications to environmental science and agriculture. Wiley, New York

Wong M (2005) Visual symptoms of plant nutrient deficiencies in nursery and landscape plants. Soil and Crop Management. SCM 10, Cooperative Extension Service, University of Hawaii, Manoa

Wrigley RJ, van de Graaff R, Lang S (2006) Residential development on the New Volcanics: a case study of suspected barium contamination and soil amendment. Austr Geomech 41(3):107–113

Yuan D, Zhang G, Gong T (2007) Variations of soil phosphorus accumulation in Nanjing, China as affected by urban development. Plant Nutr Soil Sci 170:244–249

Zhang MK, He ZL, Calvert DV, Stoffella PJ (2003) Colloidal iron oxide transport in sandy soil induced by excessive phosphorus application. Soil Sci 168(9):617–626

Zhang X, Chen J, Tan M, Sun Y (2006) Assessing the impact of urban sprawl on soil resources of Nanjing city using satellites images and digital soil databases. Catena 69:16–30

Zhang X, Tan M, Chen J, Sun Y (2005) Impact of land use change on soil resources in the peri-urban area of Suzhou city. Geogr Sci 15:71–79

Zikeli S, Jahn R, Kastler M (2002) Initial soil development in lignite ash landfills and settling ponds in Saxony-Anhalt, Germany. Plant Nutr Soil Sci 165:530–536

Part II
Site Investigation

Chapter 3
A Practical Approach for Site Investigation

Frank P.J. Lamé

Abstract Given the volume of work and costs related to soil investigations, strategies and techniques for the investigation of potentially contaminated sites have been developed and have been standardized, both on a national as well as on an international level. Investigating soil contamination is not an easy task, since contaminants are generally distributed highly heterogeneous in the soil. It is important to focus on the objectives for site investigations. In this chapter three kinds of investigation phases have been described, these are the Preliminary Investigation, the Exploratory Investigation and the Main Investigation. The Preliminary Investigation, this is the most essential phase of the whole investigation, is a desk study combined with a site visit. A Preliminary Investigation can be performed both for sites where contamination is expected and for sites that are probably uncontaminated. The main objective of the Exploratory Investigation is to proof that the assumptions made in the Preliminary Investigation are indeed correct. The goal of the Main Investigation is to provide the necessary information to deal with the contamination on a cost-efficient basis. The Main Investigation is an iterative process, where after each step the question has to be answered if the available information is 'fit for purpose'. Moreover, different sampling patterns and techniques are discussed in this chapter.

Contents

F.P.J. Lamé (✉)
Deltares, Utrecht, The Netherlands
e-mail: frank.lame@deltares.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_3,

3.1 Introduction

Historically, the investigation of soil quality had an agricultural background and was the domain of agriculturally oriented soil science. During the 1980s, soil contamination became an important issue, thus a new type of soil investigation had to be developed: the investigation of (potentially) contaminated sites. This proved to be a fundamentally different type of soil investigation compared to what was previously common practice in agricultural science. Different not only because its aim is to find contaminants in the soil rather than nutrients but, for example, also different due to the presence of buildings, roads and sealed surfaces on the site which hinder sampling on every randomly chosen location. Different also because knowledge of former (industrial) processes had to be implemented in the process of choosing where to sample, as well as for identification of the contaminants that had to be investigated.

Some decades later, the investigation of the soil quality of (potentially) contaminated sites in industrialised countries has become an area that employs far more people than the original agricultural investigations. Given the volume of work related to this type of soil investigation, both strategies and techniques for the investigation of potentially contaminated sites have been developed and have been standardized both on a national and an international level.

3.2 Not an Easy Task

Investigating soil contamination is not an easy task. Contaminants are most often distributed highly heterogeneously in the soil, due to both the way the contaminants get into the soil (e.g. industrial spilling, accidental spills or dumping) and to natural processes in the soil like groundwater transport. Moreover, we can only see the surface of the soil, while the soil itself can be highly variable, for example when small layers of sand, peat and clay alternate with each other. Variations in soil type will result in different behaviour of contaminants, showing, for example, preferential pathways in sandy soils and a high degree of retention in clay layers. The differences in interaction between various (combinations of) contaminants, the natural processes and the variations in the soil itself, altogether will result in a complex distribution of the contaminant in the soil.

Given that complexity, non-experts should be aware of the fact that there is no easy way to cope with these kinds of investigations. Although this article is partly based on standardized approaches, expert knowledge is always an important factor in the site specific design of an investigation. In the translation from the standardized approach towards a site specific approach, the investigator has to cope with the local complexity. Often there are no “easy results”.

3.3 Objectives for the Investigation of Soil Quality

Before discussing different approaches for the investigation of potentially contaminated sites, it is important to focus on objectives for site investigations. Early in the process of recognizing that soil contamination was an important issue, in the 1980s, the simple suspicion that a site might be contaminated was sufficient to start a soil investigation and subsequent remediation. And indeed, to some extent this is still a relevant motive. However, as through the years more and more contaminated sites were discovered, it became obvious that all these sites cannot be investigated and remediated within a short period of time. Therefore, it became necessary to prioritise among sites. Moreover, other motives began to play a role in the investigation of potentially contaminated sites. Today, the investigation of (potentially) contaminated sites is often initiated when sites are to be redeveloped, while other sites, although contaminated, will have to wait. Of course, when there are serious threats to humans or the environment, even derelict sites might need urgent remediation, but as long as there is no secondary incentive, it might be hard to find sufficient funding for the investigation and potential subsequent remediation of a site.

In case of secondary incentives, such as for example the redevelopment of a former industrial site as a residential area, the overall site investigation should not only focus on obtaining data on the contaminants in soil and groundwater. It should also take account of the site’s future use as well as the method of remediation that will be applied, where this is known. The investigation has to provide data relevant to the site’s future use, as well as data for the intermediate period of remediation.

Imagine a site where old building waste, containing relatively high concentrations of heavy metals, has been used to raise the level of the ground. It is obvious that contact with that layer of building waste has to be prevented if this site is to be reused as a residential area. This might be done by removing the contaminated waste. However, apart from the fact that removal of all the waste would be costly, it implies that a large quantity of new soil material would be necessary to keep the soil surface at the desired level. A better solution might be, when the contamination is immobile, to cover the contaminated building waste with clean soil and thus prevent exposure of humans to the contamination. Identifying in advance two different ways to deal with the contamination on the site in this simple example, either removing all of the building waste or covering it, will influence the type of investigation necessary. In summary:

- When removing building waste:
  - The environmental quality of the building waste has to be determined – probably an estimate of the mean concentrations of quantities of several hundred tonnes might be sufficient.

  - The total quantity of the building waste has to be estimated in order to allow costs calculations.
  - The environmental quality of the new soil material that will be used to replace the building waste has to be determined.
- When covering the building waste:
  - The immobility of the heavy metal contamination has to be evaluated.
  - The environmental quality of the new soil material that will be used as a cover layer has to be determined.

As might be clarified by this example, the early identification of potential solutions to the contamination problem is of major importance for the type of site investigation. Thus, the site investigation should not only focus on determination of the concentrations of contaminants that can be found in soil and groundwater and the spatial distribution of those contaminations. It also should deliver information that is necessary to perform a Risk Assessment, to identify potential exposure pathways, to obtain information on the geology and hydrology of the site and on future foreseeable events (e.g. flooding), etc.

Investigation of the nature and extent of (potential) soil contamination will usually be the main objective. However, that objective will usually be formulated in such general terms that further delineation will be necessary. This can be done through the identification of "subsidiary objectives" or technical goals.

3.4 Technical Goals

The identification of technical goals that are relevant for a specific site enables the investigator to define a site-specific approach for the investigation. Using a site-specific approach is of course of major importance for a cost-effective investigation. The technical goals should be formulated in such a way that these indeed provide direct guidance to the investigator. Technical goals therefore are:

- To identify the contaminants that are, or may be, present on the site.
- To determine the spatial distribution of the contaminants over the site, both in the horizontal as well as in the vertical direction, covering, when relevant, the solid, gaseous, as well as the aquatic phase of the soil.
- To determine the (potential) mobility of contaminants.
- To identify receptors (e.g. humans, the ecosystem, the groundwater) that are currently at risk, or that might be at risk after redevelopment of the site.
- To identify the pathways by which the receptors might be exposed to contaminants.
- To provide data and other information that can be used to quantify the Risk Assessment for the site.
- To provide information that will be necessary for the design of protective or remedial measures to be taken on the site.

- To determine the environmental quality of contaminated soil, or waste materials excavated from the soil, to ensure safe handling and subsequent treatment or disposal of these materials.
- To provide information for assessing legal liabilities associated with the contamination history of the site.
- To determine any immediate actions necessary to protect currently exposed receptors on the site.

This list of technical goals is just an example, but even based on this list, it can be concluded that the technical goals have a much wider meaning than solely identifying the concentration of contaminants in soil or groundwater.

Additional technical goals can and should be formulated for each site that is to be investigated. Based on these technical goals, the investigator has to determine the type and intensity of the investigation that is necessary for the site.

3.5 Three Investigation Phases

For the investigation of (potentially) contaminated sites prior to remediation, three different investigation phases are recognised:

- The Preliminary Investigation – see Section 3.6.
- The Exploratory Investigation – see Section 3.7.
- The Main Investigation – see Section 3.8.

The terminology used for these investigation phases might vary when going through the international literature on site investigations. But independent of the terminology, the distinction in three phases and the characteristics of these three phases are internationally well recognised. The main questions and actions that are addressed in these three phases are depicted in Fig. 3.1.

The Preliminary Investigation is a desk study, combined with a site visit. No samples are taken in this phase. Nevertheless, it is without doubt the most essential phase of the whole investigation. A Preliminary Investigation can be performed both for sites where contamination is expected and for sites that are probably uncontaminated. The latter sites might, for example, be future residential areas where confirmation is necessary of the fact that there is indeed no contamination present, and consequently that there are no objections from this perspective to use of the site as a residential area without further protective measures. But the investigation of sites that are expected to be uncontaminated might also be relevant for industrial sites in relation to present or future legal liabilities. As the normal exploratory investigation for industrial sites will focus on the parts thought to be contaminated, no information will be obtained for the rest of the site. To ensure that no contamination is present at the non-suspected parts of the site, an exploratory investigation of areas believed to be uncontaminated will be a valuable addition to the exploratory investigation for the suspected parts of the site.

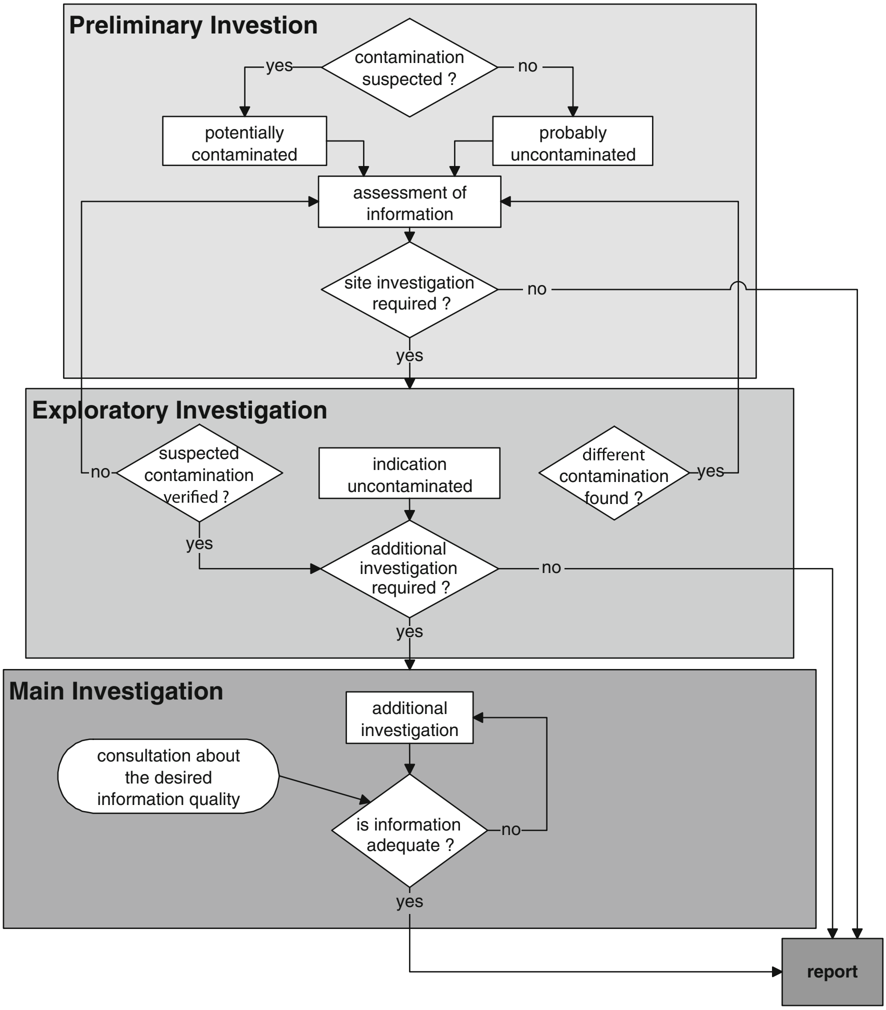


Fig. 3.1 The three phases of a site investigation and the main questions and actions for these phases

During the Preliminary Investigation, information on the site is obtained and based on that information it should be decided if there is indeed a need to sample the location. Sampling is not per definition in all situations necessary. It has to be accepted, however, that the current practice is that at least some data is necessary to provide sufficient proof for the absence of soil contamination. If however contamination is expected, the conclusion of the Preliminary Investigation will be that sampling is necessary to confirm, or deny, the findings of that investigation.

The second phase of the investigation is an Exploratory Investigation. In this phase limited sampling is performed in order to confirm or deny the expectations based on the Preliminary Investigation. The number of boreholes and monitoring wells in this phase is limited in order to obtain a cost-effective first indication of what actually is the soil and groundwater quality status of the site.

Three potential results can be obtained from the Exploratory Investigation:

- The expected contamination has indeed been found, or the expectation that there is no contamination on the site is acknowledged; in each case the assumption of the soil and groundwater quality of the site is confirmed in the Exploratory investigation.
- The expected contamination has not been found and thus the assumption of the soil and groundwater quality is denied.
- A different, not expected, contamination is found.

Clearly all three options will have quite different consequences for the conclusions of the Exploratory Investigation and the potential further steps, which will be discussed in more detail in Section 3.7. In general, however, if a contamination has been found, it is necessary to obtain more information on that contamination, and fulfilling the technical goals as described in Section 3.4 becomes necessary. This implies that a Main Investigation will be necessary. Obviously, part of the focus of the Main Investigation is to obtain detailed information on the spatial distribution of the contaminants. However, to what extent there is a need for that detailed information will depend on the Risk Management solution, e.g. the remedial action, that will be performed. See also the example in Section 3.3. So the Main Investigation will be an iterative process, where after each step, the question has to be answered if the available information is "fit for purpose".

These three phases will be discussed in more detail in the next sections.

3.6 Preliminary Investigation

As mentioned in Section 3.5, the Preliminary Investigation is in fact the most important phase of a site investigation. When no or insufficient information for the site is obtained in this phase, it is highly likely that samples obtained during the Exploratory Investigation will be taken at wrong locations and consequently might provide a highly biased impression of the soil quality! Sites that are contaminated with highly toxic contaminants might be considered uncontaminated due to the large spatial variation often observed in soil. Assumptions on the type and spatial distribution of the expected contaminants are therefore of major importance as it will be neither financially nor technically possible to define an effective investigation strategy without assumptions.

Assumptions on the potential contamination of the soil and groundwater are based on knowledge of the (former) activities and processes on the site and are to be combined with knowledge of the local geological and hydrological situation. Figure 3.2 shows the gas purification building of a former gasworks site in

Fig. 3.2 Gas purification building of the former gasworks in Rotterdam Kralingen (built between 1884 and 1887) with iron bearing soils ("bog ore"), used for trapping of cyanide in raw coal gas, laying outside for regeneration

Rotterdam (Kralingen) in the Netherlands, built between 1884 and 1887. In order to regenerate the iron bearing soil material, the so-called bog ore, which was used to trap cyanide from the coal gas (formation of "Berlin Blue" or "Prussian Blue"), the bog ore was spread outside in a thin layer. Obviously, this has had severe consequences on the soil quality of the site.

Additionally, a photo from the same gasworks site (Fig. 3.3) shows two employees in the late 19th century using a simple pump to extract something from

Fig. 3.3 Men at work at the former gasworks in Rotterdam Kralingen. What were they pumping?

underneath a road on the gasworks site. Considering their clothing it might well have been oil or tar that they were pumping. However, the essential message from a picture like this is that it shows an activity on an unexpected location which was, probably more or less accidentally, caught on photo. It implies that, whatever effort will be put into the Preliminary Investigation, one has to accept the fact that not all potentially contaminating activities can be traced. Certainly not for a site that has been used for several decades or even centuries. This lack of knowledge of past activities may not be an excuse to omit the Preliminary Investigation. The example shows that soil contamination can be present due to different (often unknown) activities and consequently that sampling only very small locations on a site, per definition implies a risk of missing the contamination.

Let us assume another example where a drum containing dioxins is buried underground in an area of 50 by 50 m. This is only a small area which has been identified as the approximate location where the drum was ditched in the past. Getting a more precise indication of its location is not very likely in practice. The surface of the area is 2,500 m^2, while the surface of the drum is assumed to be 1 m^2. Consequently, using one boring in the Exploratory Investigation to determine if there is a contamination, would result only in a 1 to 2,500 chance (0.04%) of actually finding the drum! Or enhancing the probability of finding the drum to 90% would mean that 2,250 borings are necessary. Obviously, where in the Preliminary Investigation only limited sampling will be performed, the single boring and thus the probability of 0.04% is more in the direction of what is to be considered as an acceptable effort than the 2,250 borings.

These examples show the importance of the Preliminary Investigation based on which it will be determined where samples are to be taken. They also illustrate the limitations of the conventional technique for soil investigations, i.e. using (hand) augers to make borings and taking samples to the laboratory. Other survey techniques might be far more successful in finding the drum, even without actual intrusion into the soil. A number of different non-intrusive techniques can be mentioned, e.g. magnetometry, metal detection, ground penetrating radar, or seismic/acoustical techniques. However, although in recent years large enhancements in the performance of the previously mentioned techniques have been achieved, specifically by combining several techniques, the circumstances in the soil will determine the success of these techniques to a large degree. When for instance the soil contains a lot of debris, when the groundwater table is near the surface, or when the soil is a heavy clay, the performance of these techniques is limited. Nevertheless, also the “traditional” technique of borings has under specific circumstances, as shown in the example, only low performance characteristics. Better performance might be obtained by excavating trial pits with a mini excavator, but at the same time the potential exposure risks are much larger compared to borings. However, specifically in soils with more coarse material, the excavation of trial pits can be far more practical than borings.

Apart from identifying the locations where a contamination might be present, the Preliminary Investigation should therefore also provide indications on the most optimal survey technique(s) for that specific location in the Exploratory Investigation.

During the Preliminary Investigation all information necessary to obtain a first indication of the expected contamination of the site is to be gathered. However, the complexity of the history of the site, its use, as well as the complexity of the local geological and hydrological situation, will determine the necessary effort. Additionally, the future use of the site plays a role in the outline of the Preliminary Investigation and so does the potential necessity to obtain data for legal liabilities. Therefore the desirable level of detail to be reached in the Preliminary Investigation might vary from site to site, but can be defined in a limited number of classes to obtain a more standardized approach. For each of these classes, information is to be obtained on the following issues:

- Former use of the site:
  - Information can be obtained from old maps, former employees, photos, aerial photos, process knowledge of (old) manufacturing processes, archives, reports of earlier site investigations, et cetera.
- Current use of the site:
  - Information can be obtained from a site visit, interviews with current employees, neighbours and site owner, et cetera.
- Future use of the site:
  - Information can be obtained from the site owner, local or regional authorities, et cetera.
- Local geology, soil type and soil characteristics;
  - Information can be obtained from the national geological survey, soil maps, reports of earlier site investigations on the site as well as near the site, et cetera.
- Local hydrology:
  - Information can be obtained from the national geological survey, reports of earlier site investigations on the site as well as near the site, et cetera.
- Financial and juridical aspects:
  - Information can be obtained from the owner of the site, owner of the buildings on the site, et cetera.

Obtaining historical information of a site is not only relevant for a specific period. Often the developments on a site will provide valuable information on the site's history and the activities that took place over the years. When available, aerial photos can provide great support, even by coincidentally catching the act of soil contamination, but surely provides information on the site's history over a period of years when more than one aerial reconnaissance is available.

The purpose of the Preliminary Investigation is that the investigator obtains a first impression of the site and the contamination that might be present. At first, this impression will only be a rough "sketch" of potential contaminated locations on a local geographical map. The Preliminary Investigation probably results in a vague indication of the area that might potentially be affected by the contaminating processes that took place on the site.

If more precise information is available, a combination of old and new maps is an interesting option as, by this combination of maps, potentially affected locations can be pinpointed more precisely. This is specifically favourable if the use of the site has changed considerably over the years, for example where a former industrial site has already been used as a residential area for a number of years.

The first impression of the investigator is known as the "conceptual model". As said, at first the conceptual model will be no more than a sketch and the term might indicate a much higher level of detail than is actually reached. However, it is the conceptual model of the site and its contamination that will be maintained, enhanced and adjusted over the various phases of the investigation, until it has reached such a degree of detail that it is sufficient in light of the Risk Management activities that will be performed on the site.

To establish a first conceptual model based on the Preliminary Investigation, a number of questions are to be answered. These questions are, for example:

- What processes happened on the site?
- On which exact locations did these processes take place?
- Which contaminants were involved in these processes?
- Through which mechanism might these processes have had an adverse effect on the soil quality?
- Were there protective measures on the site at the time when these processes took place?
- What is the effect of the local geological situation as well as the local hydrological situation in relation to the potential spatial distribution of contaminants (occurring transport processes)?
- Where (horizontal and vertical) is the most likely location of the contamination?
- Is the contamination only expected to be present in the soil, or also in the groundwater?

The investigator should bear in mind that soil contamination is not per definition caused by accidental spillage or leakage. For former industrial processes, or for current industrial processes in countries where the environmental legislation is yet insufficiently developed, deliberate dumping of contaminated substances in, for instance, soil or sewage water, was and is an accepted process of dealing with waste materials. These (formally) accepted ways of dealing with wastes, might have (had) an important adverse effect on soil quality.

Performing a Preliminary Investigation implies that one has to visit the site. A site reconnaissance is considered an essential step in the Preliminary Investigation as the investigator can experience the current situation at the site. This is also relevant for the later Exploratory and Main Investigation, for example to investigate where samples can be taken and where not. The site reconnaissance provides, in addition to already available maps and/or photo's, essential information on the current state of the site and its buildings. Often this provides important clues to where contaminating activities might have taken place. See for example Fig. 3.4 where dumped

Fig. 3.4 Dumped drums in the backyard of a metal workshop

drums were found in the back of an industrial site and Fig. 3.5 where storage facilities are so poor that the contamination of the soil surface is already visible during the site reconnaissance. Obviously, observations like these will result in sampling locations for the Exploratory Investigation. Sampling locations that would not have been selected in the sampling plan would be based on a Preliminary Investigation without a site reconnaissance.

As the Preliminary Investigation, apart from a visit to the site, only consists of studying archives, it is hard to assure that an investigator indeed has put sufficient

Fig. 3.5 Storage of drums without proper soil sealing, resulting in soil contamination due to accidental spills

effort in obtaining information about the site. It is quite easy to state that "no information has been found" as checking if that information is indeed not available would imply that at least part of the Preliminary Investigation has to be performed again. Consequently, there is a tendency in a competitive market to limit the effort that is put into the Preliminary Investigation. Two potential mechanisms are important to ensure the quality of the Preliminary Investigation, i.e. detailed instructions on the activities the investigator has to perform (1) and quality control through, for example, certification of the investigator (2). The combination of these two instruments, a standard with a detailed technical description and a certification scheme, at least provides a mechanism of quality control for the Preliminary Investigation. An additional control instrument is the obligation to publish a transparent report on the Preliminary Investigation.

The report of the Preliminary Investigation must contain both factual information as well as interpreted information. Examples of *factual* information are:

- Old maps, showing the location of specific buildings and the boundaries of the site.
- Old photos of the site.
- Registrations of interviews with former employees (although the information provided is per definition an interpretation of the person interviewed).

In short, the factual information is all information as obtained during the Preliminary Investigation. It is important to include the factual information in the report of the Preliminary Investigation, as through this route the information is still available, even if another investigator or other consultancy firm takes over the investigation of the site. Repetitive misinterpretations of the information can be prevented by providing the factual information in the report.

Interpreted information relates to that part of the report in which the investigator, based on the obtained factual information, describes what to his opinion might have happened on the site and what the consequences thereof are for the soil quality. In fact, this interpretation provides the first conceptual model of the contamination.

In order to provide guidance to the investigator, national and international standards for the performance of the Preliminary Investigation use a limited number of hypotheses for the description of the spatial distribution of the contamination. A, to some degree layered, approach is used where a distinction between sites at first is based on:

- No contamination.
- Contamination.

As a secondary assumption, for the contaminated sites a distinction is made between:

- Local soil contamination.
- Diffuse soil contamination.

Local soil contamination implies that a specific process is expected to have had an adverse affect on a specific location. This was not necessarily a small area, but obviously, when the area gets bigger the contamination within that area will appear more or less as diffuse contamination. The practical distinction between these two types of contamination therefore is that on the scale of the investigation, only a part of the site is contaminated when it is considered to be local soil contamination, while the whole site and more will be contaminated when it is diffuse soil contamination. But also the contaminating process will determine if the contamination is to be considered local or diffuse.

Finally, a third assumption deals with the more detailed expectations of where the contamination can be found. For local soil contamination, the contamination has either a:

- Known position of the core of the contamination.
- Unknown or unsure position of the core of the contamination.

For a diffuse contamination, the contamination is either:

- Homogeneously distributed.
- Heterogeneously distributed.
- The boundaries of the contamination are unknown or uncertain.

As a consequence, there are different types of diffuse soil contamination. A contaminating source at some distance, for example stack emissions, might result in a contamination in the top soil where little variation is expected between individual locations on the site. The contamination, originating from a diffuse source, can be considered as homogeneously distributed. But when harbour sludge is used in embankments, large differences on a small scale can be expected. However, that same expectation on variability and contaminants applies for every location in that embankment and therefore the embankment as such can be considered as diffuse, but heterogeneous, contamination. Finally, diffuse contamination, being either homogeneous or heterogeneous, might have boundaries that are relevant on the scale of the investigation. In these situations, finding the boundary of the contamination is in principle more relevant than investigating the variation in concentrations within the contaminated spot.

The Preliminary Investigation ends with the assumptions on the presence and spatial distribution of the contaminants in soil and/or groundwater. These assumptions are formalized in national and international standards in hypotheses which are used as a basis for starting the Exploratory Investigation.

3.7 Exploratory Investigation

The Exploratory Investigation is, after the site visit during the Preliminary Investigation, the second time that the site is actually visited. However, it is the first time that samples will be taken. Prior to sampling, first a sampling strategy has to be defined. Obviously, the sampling strategy is based on the results and the resulting hypotheses of the Preliminary Investigation.

The main objective of the Exploratory Investigation is to prove that the assumptions made in the Preliminary Investigation are indeed correct. In science a hypothesis is only accepted when it cannot be rejected. For soil investigations we, on the contrary, assume that the hypothesis is correct when the results are, more or less, in line with what was expected. This once again stresses the importance of a good Preliminary Investigation, since there is only a small chance of rejecting the hypotheses of the Preliminary Investigation based on the results of the Exploratory Investigation.

Additionally, the amount of effort, and thus money, that we want to put into the Exploratory Investigation is limited. Theoretically, the optimal balance should be found between the amount of money that is needed to conduct the Exploratory Investigation and the (unknown) risk of not finding an unknown contamination. This is per definition an “unsolvable equation”, so the effort for an Exploratory Investigation is based on the opinion of experts. On a national level in the Netherlands, there is agreement on the level of detail, on the number of borings, samples and monitoring wells for many individually distinguished types of soil contamination. On an international level such agreement is hard to reach, as the boundary conditions will vary between countries. If, for example, there are hardly any specialized environmental analytical laboratories in a country, prices of analyses might be significantly higher than for a country where a number of laboratories operate in a competitive market. Similar arguments can be used for consultancy firms performing the investigations. At the same time, the perception of what is acceptable and what is not, will also vary between countries. Consequently, it does not make sense to define the number of borings, samples and monitoring wells in an international standard.

As mentioned in Section 3.6, it should also be kept in mind that borings are not always the most reliable, or even practical, way of obtaining data from a site. Per definition the obtained quantity of soil material is very limited with respect to the soil volume that is to be characterised, while the very nature of the soil can also set limits to the application of soil augers. Other techniques, either being intrusive of non-intrusive, can therefore be a preferred alternative.

Confirming the results of the Preliminary Investigation implies that we want to sample the soil and groundwater at each location within a site where soil pollution might be present. And in order to enhance the chance of success, samples of these locations will be taken where the highest concentrations (the “hot spot”) might be expected. In principle, a single sample with a high concentration proves that the expected contamination is indeed present. Subsequently, an assumption might be made on the necessity to perform a Main Investigation, in conjunction with the results of the Preliminary Investigation. From this perspective, the Exploratory Investigation is only a step of minor importance between the Preliminary Investigation and the Main Investigation. And to some degree that is true, specifically for situations where an obvious serious contamination in soil and/or groundwater is present. However, for less obvious situations, and specifically when no contamination is found, the Exploratory Investigation will be the one and only investigation phase where actual samples are taken and analysed. The Exploratory Investigation can also be the only investigation phase in situations where a decision is to be made about whether to buy a site.

Thus, the Exploratory Investigation should be sufficiently intense to be able to conclude that further investigation of the site is not necessary. So again from this perspective, defining the effort to be put in the Exploratory Investigation is finding a balance between the costs for the investigation and the quality and acceptable "level of proof" provided by the Exploratory Investigation.

The fieldwork for the Exploratory Investigation is based on the hypotheses as defined in the Preliminary Investigation. In simple situations a single hypothesis describes the expected contamination and can be used to investigate the site. An example is a site where there is an underground storage tank for mineral oil present in the soil and the objective of the investigation is to determine if that tank is leaking or has contaminated the soil. The hypothesis in this particular situation is that there is a local soil contamination with a known location. A limited number of borings, covering both the direct surrounding of the tank, the pipes connected to it and the feeding pipe, will provide sufficient evidence for acceptance or rejection of the hypothesis.

In more complex situations, combinations of various types of contamination might be present. Contaminations with different type of contaminants, different processes that caused the contamination and differences in spatial distribution of the contaminants. Additionally, there might also be a need to obtain a more general impression of the soil quality of the whole site, a goal that can also be part of the Exploratory Investigation. Obviously, these goals are not to be dealt with separately. A combination of sampling strategies for the individual hypotheses to be tested, can result in a more cost-effective approach for the investigation, without losing quality for each individual hypothesis.

As discussed previously, the number of borings, samples and monitoring wells as mentioned in the Dutch standard, cannot be regarded as the true and only answer. At the same time, these strategies can be seen as exemplary for what can be used for the different hypotheses, in every part of the world.

The Dutch standard for the Exploratory Investigation identifies the sampling strategies for the following types of sites:

- Unsuspected site.
- Large scale unsuspected site.
- Suspected site with local soil contamination with known location.
- Suspected site with one or more underground storage tanks.
- Suspected site with homogeneously distributed diffuse soil contamination.
- Suspected site with heterogeneously distributed diffuse soil contamination.
- Suspected site with unknown soil contamination.
- A baseline investigation for future potential contaminating activities (defining the soil quality prior to the soil use that might have an adverse affect).
- A baseline investigation for future underground storage tanks.

In addition to the investigation of the sites, the standard also contains three strategies which aim at the determination of the soil quality of soils to be excavated in the future:

- The determination of the soil quality for uncontaminated soils.
- The determination of the soil quality for large scale uncontaminated sites.
- The determination of the soil quality for soils with a heterogeneously distributed diffuse soil contamination.

The latter three strategies are to some degree comparable to the site investigations for these types of sites, but comply with legislative boundary conditions set for the reuse of soil lots in the Netherlands.

For each strategy, concrete instructions are formulated on:

- The sampling pattern.
- The maximum depth of borings.
- The soil layers that are to be sampled.
- The number of borings for soil sampling.
- The number of borings where monitoring wells are to be installed.
- The number of samples of both soil and groundwater that actually are to be analyzed.
- If soil samples are to be analyzed on an individual basis, or that the analyses are performed on composite samples.
- The contaminants that are to be analysed in soil and groundwater samples.

Two examples of these sampling strategies are provided in Tables 3.1 and 3.2. It is important to realize that the numbers provided in these tables do not have a scientific basis. These are, as described before, thought to be a good compromise between the costs of sampling and analysis in comparison to the type of contamination to be investigated and the potential risk of missing a contamination.

As mentioned before, these strategies are based on the application of conventional sampling techniques, thus by obtaining samples through the application of

Table 3.1 Specifications for the investigation of a suspected site with local soil contamination with known location

	Number of borings		Number of samples to analyse	
Surface (ha)	To 0.5 m under contamination	Boring with observation well	Soil from hot spot	Groundwater
<0.01	2	1	1	1
$0.01 \leq 0.05$	3	1	1	1
$0.05 \leq 0.08$	4	1	1	1
$0.08 \leq 0.10$	5	1	1	1

Table 3.2 Specifications for the investigation of a suspected site with homogeneously distributed diffuse soil contamination

	Number of borings			Number of (composite) samples to analyse	
Surface (ha)	Down to 2 m	With observation well	Total	Soil (suspected depth)	Groundwater
<1	5	2	7	2	2
1	6	3	9	3	3
2	7	3	10	3	3
3	8	4	12	4	4
4	9	4	13	4	4
5	10	5	15	5	5
6	11	5	16	5	5
P	$5 + p$	$2 + 0.5p$	$7 + 1.5p$	$2 + 0.5p$	$2 + 0.5p$

(hand driven) augers, the selection of samples in the field and sending these samples to the laboratory for chemical analyses. Consequently, they are appropriate when there are no possibilities for the application of other techniques (e.g. magnetometry, metal detection, ground penetrating radar, seismic/acoustical techniques, trial pits). However, application of those techniques often can deliver much more information for similar costs that would otherwise have been spent on sampling and analyses. Still, non-intrusive techniques and screening techniques are often less sensitive as well as less quantitative. Consequently, these techniques are of much more value when applied in conjunction with the more traditional method of sampling and analysis. Indeed, in the development of international standards on the application of screening techniques for the investigation of soil contamination, the parallel use of traditional techniques to allow quantification of the results of the screening technique, is considered as a boundary condition. Nevertheless, the number of traditional samples and analyses can be diminished to a large extent while at the same time screening methods provide the possibility of obtaining a much more dense network of measurements and consequently a much better perception of the actual soil quality. For the phase of the Exploratory Investigation, the application of screening techniques, however, is less obvious, as the purpose of the Exploratory Investigation is not to map the boundaries of the contamination, but to acknowledge that an assumed contamination is indeed present.

Having performed the fieldwork and the analyses, the results of both activities have to be interpreted in light of the hypothesis that was the basis of the applied strategy. Often there is a tendency to accept the analytical results as being a better representative result than the observations of the sampler. For example, when the sampler has observed an "oil like smell", it is often thought to be an incorrect observation when the analysis for mineral oil does not prove that there is indeed a contamination with mineral oil. In this case it is a familiar mistake to forget that the analysis for mineral oil includes oil components with an equivalent boiling point of more than C_{10}, while the more volatile components might be present, but are not detected in the laboratory due to the fact that they might have been lost from the sample (evaporation, biodegradation) or in the process of obtaining an analytical sample (evaporation) prior to analysis. So when there is a clear indication of a misfit between the observations in the field and the analysis, there might be a reason for additional sampling and/or analyses.

The power of observations in the field by the sampler, like the previously described “oil like smell”, has of course its limitations. Although the nose is a very sensitive “detecting instrument”, it is also easy to obtain highly biased “measurements”. The sensitivity of the nose is affected by the overall smell in the area, becomes less sensitive after longer exposure periods and, moreover, the sensitivity will be very low when the sampler has a cold. Nevertheless, field observations are important and have to be reported by the sampler. More important, however, is the fact that smelling will expose the sampler to contaminants that might be highly toxic. For this reason, organoleptic observations, apart from visual observations, are to be prevented. Here again, there might be a role for screening techniques. Although these techniques are less easy to operate than the nose, they are more quantitative and objective and therefore easier for interpretation by the consultant who has to write the report on the Exploratory Investigation.

When writing the report of the Exploratory Investigation, it is important to include the observations of the sampler, the soil descriptions made and any indication of potential contaminations as observed in the field. Obviously, also the analytical results are to be included in the report. All information mentioned previously can be considered as factual information. But at least as important is the interpretation of that information in an enhanced and perhaps adjusted description of the conceptual model of the site and its contamination. As mentioned for the report of the Preliminary Investigation, the consultant has to make a clear distinction in the report on the Exploratory Investigation between facts and the interpretation based on that.

The hypothesis which was used to define the sampling strategy is to be tested in light of the obtained results. Do the results indeed acknowledge the conclusions of the Preliminary Investigation, or is there something different on the site? And if there appears to be a difference, what does that imply for the further investigation of the site? Are the results conclusive enough to proceed to the Main Investigation phase, or is an additional phase of the Exploratory Investigation necessary to increase the reliability about the type and spatial distribution of the soil contamination? Obviously, the latter can also be part of a first step of the Main Investigation. However, the main purpose of the Main Investigation is to delineate the contamination that was identified during the Preliminary Investigation and of which the presence was affirmed during the Exploratory Investigation. Contaminations that are not considered in these two phases, either due to a lack of information in the Preliminary Investigation or due to missing the correct location in the Exploratory Investigation, have an enlarged possibility of not being detected in the Main Investigation. So before going into the phase of the Main Investigation, it is important to be as sure as possible that all potential contaminations on the site are identified.

So in addition to reporting factual information obtained in the Exploratory Investigation, the report on the Exploratory Investigation also has to discuss the hypotheses that were on the basis of the investigation and define if the obtained results confirm or deny them.

Finally, the question if a Main Investigation is necessary has to be answered during the Exploratory Investigation. Obviously, when no contamination was found and this conclusion is in line with the hypothesis (uncontaminated site), there is no need for conducting a Main Investigation. This can be the case, for example, when the location of an underground storage tank has been well established during the Exploratory Investigation, but when no contamination has been found. The obvious conclusion is that there is no contamination originating from the tank and consequently the results of the Exploratory Investigation are sufficient to end the investigation of the site.

But even if the hypothesis has proved to be incorrect, the need for a Main Investigation might be absent, for example when the concentrations found are enhanced in comparison to local background values and/or Soil Quality Standards, but still are relatively low so that there is no apparent need for further actions. In this case we also have to be careful that accidentally the samples in the Exploratory Investigation were taken on the boundary of the contamination and not in the centre, thus implying lower concentrations than are actually present on the site.

Finally, no additional investigation might be necessary when the objective of the Exploratory Investigation did not originate primarily from an environmental perspective, such as for example when it was performed in relation to the decision whether to buy a site.

In all other situations, thus where the contamination was actually found (in sufficiently high concentrations), there will be a need for a Main Investigation.

3.8 Main Investigation

The goal of the Main Investigation is to provide the necessary information to deal with the contamination, the presence of which was confirmed during the Exploratory Investigation. What level of information is necessary, both for the contamination as well as for the site characteristics, will depend on a variety of site specific parameters, but the aim is to perform a cost-effective investigation.

The Main Investigation should be the final investigation phase prior to management actions (e.g. remediation). That implies that, apart from obtaining a more detailed view of the spatial distribution, the Main Investigation should also provide information that can be used to determine the urgency of a remediation, in regard to the present or future land-use, as well as information that is essential for the application of remedial actions. Consequently, the Main Investigation should not only focus on the contamination, but should also take account of quite different factors, like the identification of potential exposure pathways. When there is clarity on potential remediation measures, the Main Investigation should take account of the information necessary to enable specific types of remediation. When for example in-situ bio-remediation is a possible and preferable option, the Main Investigation should also provide information on the presence of bacteria in the soil that can be stimulated and their need for nutrients and/or oxygen for enhanced biodegradation of the contaminants. However, it is not always possible to predict the potential remedial techniques in advance, while there might also be practical, temporal or legal reasons not to focus (too much) on the information necessary for the remediation. In those cases the Main Investigation has a focus on the contamination.

More in general, and independent of the potential techniques that can be applied to remediate the site, knowledge of the local geology, soil type and hydrology are always of major importance. Not only because these characteristics will have an important influence on the spatial distribution of the contamination, but also because this information is essential for a good judgment on subsequent steps, both in light of the continuation of the investigation, if necessary, as well as for later Risk Management actions (e.g. remediation).

In general, the Main Investigation is, for the part of the investigation that focuses on the contamination, highly comparable to the Exploratory Investigation. However, instead of seeking confirmation of the presence of the contamination, the Main Investigation should focus on finding the boundaries of the contamination in both the horizontal and vertical directions. And apart from investigation of contamination in the soil, the volume of groundwater that is contaminated must be determined. Apart from defining the delineation, it is also relevant to obtain information on the concentrations that can actually be found in the "hot spots", as these might also determine possible options for dealing with the contamination.

Finding the boundaries of the contamination and at the same time obtaining additional information necessary for potential remedial actions, will often need a number of subsequent steps in the Main Investigation. The number of steps necessary depends, of course, on the size of the contamination, the distance between sampling points in the various phases, the complexity of the contamination on the site, the level of detail needed in light of potential remediation techniques, as well as the use of additional information, like the results of field screening methods. Consequently there cannot be a simple guideline on how the goal of the Main Investigation can be reached as efficiently as possible.

Obviously, after each step, the investigator has to take account of the obtained results when planning additional sampling. More specifically, the obtained results are to be compared with the hypothesis (or a combination of hypotheses) on the contamination that forms the basis of the investigation. If in a specific step the results deviate from the hypothesis, a definition of an alternative hypothesis might be necessary, which would result in a better guidance for additional sampling steps. Again, the consultant should make a clear distinction between facts (e.g. analytical results, field observations) and interpretations (e.g. which samples are actually analysed, what is concluded on the obtained results).

After each step of the Main Investigation, the conceptual model of the contamination and the site becomes more detailed. Still, even when a large number of samples are taken, one has to realise that only a very small portion of the soil has been investigated. Consequently, even with the most intensive Main Investigation, one has to accept the fact that there will appear differences between the conceptual model and reality. As mentioned earlier, certainly in the Main Investigation, non-intrusive techniques and screening techniques can play in important role. These techniques can, at least partly, bridge the gap between the per definition scarce information from sampling and analysis and the contamination that is really there.

3.9 Sampling Patterns

Specifically in the Main Investigation the issue of sampling patterns arises. Whereas in the Exploratory Investigation the sampling is the primary target, more structured sampling patterns can be used in the Main Investigation.

Targeted sampling, also known as judgmental sampling, is a type of sampling where samples are taken at points that are of specific interest. In the Exploratory Investigation, these are the locations where soil contamination is expected. It is obvious that the results obtained with these samples are not representative of the whole site under consideration. From a statistical point of view, these results will be highly biased if they are used to predict the quality of the whole site. Still, from the perspective of the Exploratory Investigation these results are highly valuable as they provide, based on the Preliminary Investigation, a cost effective indication of the presence or absence of contamination.

As stated in Section 3.8, the Main Investigation aims to delineate the contamination while simultaneously obtaining more information on the level of contamination of the "hot spots". Obviously, these two goals can be achieved at the same time. Consequently there is a need to sample both on the (expected) boundaries of the contamination as well as in the "hot spots" themselves. The use of a predefined sampling pattern can then be of help.

For easy interpretation, systematic sampling patterns, sampling grids, are easy to handle. However, in practice a lot of grid points will often prove to be inaccessible for sampling due to buildings on the site, underground constructions and pipeline, et cetera. Systematic patterns imply a risk of biased sampling when the sampling pattern coincides with a (historical) pattern in the site. A simple example is a pipeline in parallel with the grid lines. Leakage from the pipeline will be found when the grid is close to the pipeline, while it will never be found when the grid is too far away. In practice, these risks are not too important and can be acceptable, specifically in light of the limited soil volume that is sampled and the risk of the resulting bias. More important perhaps is the fact that a systematic pattern tends to identify a lot of samples on locations that are of less interest, as there is no real expectancy of contamination. This can be overcome by using a systematic pattern that focuses on locations of interest.

Apart from systematic sampling, other patterns are random sampling or stratified random sampling. In random sampling, the locations are chosen randomly over the site. In practice this technique will not be used very often as there is no guidance towards parts of the site that are of more interest and consequently should be sampled more intensively. The big advantage of random sampling is the fact that it provides a sound basis for statistical conclusions, for example on the mean quality of the site. Random sampling might have a field of application when determining the quality of a certain predefined quantity of soil, either being a soil lot on the surface after excavation, or still in-situ.

Stratified random sampling is far more applied than random sampling as, due to the stratification, it is impossible that all samples are by chance taken in a very small part of the site. In stratified random sampling the site is first stratified. In this stratification, one can take account of the expectation on the level of contamination. Within the strata, one or more samples are taken at randomly defined locations. The statistical value is still comparable to true random sampling as long as the strata all have the same size and in each stratum the same number of samples is taken. When locations that are of special interest have smaller strata, sampling is intensified in these areas, even when the number of sampling points within all strata is kept equal. From a statistical point of view this will result in over-sampling of certain parts of the site and consequently the determination of the overall quality of the site will than be biased. At the same time, in these investigations it is most often not the mean value of a site that is of interest.

The data obtained through sampling can be used to predict where additional sampling is necessary. This can be done in a judgemental way, where the consultant, based on his expert opinion and in light of all information available (numerical as well as non-numerical), defines new sampling locations. This is the currently most used approach. The disadvantage of this approach is that the bias cannot be defined. Sampling, and consequently the results of the investigation, is as good or bad as the assumptions made by the investigator. In practice we will be confronted with the bias when, during a remediation, the actual quantity of soil that is to be excavated is far larger than was predicted. However, there is no guarantee that a full systematic or statistical approach will not provide that risk. As long as we cannot look into the soil and contamination is distributed heterogeneously in that soil, any form of non-intensive sampling will result in bias.

Alternatively, (geo)statistical sampling techniques can also be applied, where additional sampling is to be performed where the uncertainty in predicted concentrations is too large. There have been numerous attempts to use geostatistical techniques like kriging for optimization of sampling in the Main Investigation, but in most situations the spatial relation between data is only limited when compared to the random variation in contaminated soils.

3.10 Sampling Techniques

Previous Sections discussed the strategic aspects of the investigation of a (potentially) contaminated site with most emphasis on the sampling of soil. Sampling of other media is of importance as well. During the Exploratory Investigation sampling will in most cases be limited to soil, while groundwater will be sampled when sufficiently close to the surface (for example within 4 or 5 m). During the Main Investigation sampling of groundwater is as important as sampling of soil, but obviously is also related to whether there is groundwater within any proximity of the contaminants as found in the soil and whether these contaminants are mobile. Sampling of soil gas obviously is relevant when there is a considerable unsaturated zone, or when there is a contamination in the soil or groundwater close to the surface. Sampling of soil gas is of major importance when there are buildings on the site into which contaminants can evaporate.

In addition to the choices about where and what to sample, also the question "how to sample" is to be answered. Previously, it was already mentioned that the investigation of contaminated sites still predominantly uses more traditional techniques of sampling, taking samples from the soil and groundwater and sending these to an environmental laboratory for analysis.

Descriptions of sampling apparatus and instructions on how to apply these are available in national and international standards. This is essential, as not only the applied sampling strategy will determine the representativeness, and therefore the quality, of the samples, it is also determined by correct appliance of sampling apparatus and sampling techniques.

Apart from actual sampling, other techniques are also available for obtaining information on the contamination as well as on the geology and hydrology of the site. These can be intrusive as well as non-intrusive techniques. Application of these techniques, in addition to more "classical" sampling and analysis, can result in a substantial enhancement of the image (the conceptual model) and understanding of the local situation, allowing better informed decisions than would be possible when only "classical" sampling and analysis are used.

Literature

ISO 10381-1	2003	en	Soil quality – Sampling – Part 1: Guidance on the design of sampling programmes
ISO 10381-2	2002	en	Soil quality – Sampling – Part 2: Guidance on sampling techniques
ISO 10381-3	2001	en	Soil quality – Sampling – Part 3: Guidance on safety
ISO 10381-4	2003	en	Soil quality – Sampling – Part 4: Guidance on the procedure for investigation of natural, near-natural and cultivated sites
ISO 10381-5	2005	en	Soil quality – Sampling – Part 5: Guidance on the procedure for the investigation of urban and industrial sites with regard to soil contamination
ISO 10381-6	1993	en	Soil quality – Sampling – Part 6: Guidance on the collection, handling and storage of soil for the assessment of aerobic microbial processes in the laboratory

ISO 10381-7	2005	en	Soil quality – Sampling – Part 7: Guidance on sampling of soil gas
ISO 10381-8	2007	en	Soil quality – Sampling – Part 8: Guidance on sampling of stockpiles
NEN 5706	2003	nl	Soil – Guidance on the description of sensory perceptions for environmental soil investigation
NEN 5720	2008	nl	Soil – Sediment – Investigation for exploratory survey
NEN 5725	2008	nl	Soil – Guidance on the preliminary study for preliminary site investigation, first investigation and main investigation
NEN 5740	2008	nl	Soil – Investigation strategy for exploratory survey – Investigation of the environmental quality of soil and soil lots
NPR 5741	2003	nl	Soil – Methods of boring and sampling apparatus for soil, sediments and groundwater that are applied for investigation of soil contamination
NEN 5742	2001	nl	Soil – Sampling of soil and sediments for the determination of metals, inorganic compounds, semi-volatile organic compounds and physico-chemical soil characteristics
NEN 5743	1995	nl	Soil – Sampling of soil and sediments for the determination of volatile compounds
NEN 5744	2008	nl	Soil quality – Sampling of groundwater
NEN 5766	2003	nl	Soil – Installation of wells for environmental research
NEN 5861	1999	nl	Environment – Procedure for the transfer of samples

Chapter 4
Statistical Sampling Strategies for Survey of Soil Contamination

Dick J. Brus

Abstract This chapter reviews methods for selecting sampling locations in contaminated soils for three situations. In the first situation a *global* estimate of the soil contamination in an area is required. The result of the surey is a number or a series of numbers per contaminant, e.g. the estimated mean concentration, median, 90th percentile, or the cumulative frequency distribution for the area as a whole. In the second case we want more spatial detail, and interest is in the mean or median concentration for several delineated blocks. Finally, in the third case the aim is to construct a high resolution map of the concentrations, for instance by geostatistical interpolation. For the first aim, design-based sampling methods, in which locations are selected by probability sampling, are most appropriate. Several basic sampling designs are described. Laboratory costs can be saved by bulking soil samples. The precision of estimates can be increased by exploiting ancillary information on variables correlated with the contaminants. For mapping purposes, model-based sampling methods, in which locations typically are selected by purposive sampling, are the best option. Examples are sampling on a centred grid, spatial coverage sampling, and geostatistical sampling. A simple method, based on the k-means clustering algorithm, is described for computing spatial coverage samples. For geostatistical interpolation a variogram is required. Variogram estimation is enhanced by adding several tens of locations within short distance of the locations of a grid or spatial coverage sample. A separate section describes sampling methods for detecting and for delineating hot spots.

Contents

D.J. Brus (✉)
Wageningen University Research Centre, Alterra, Wageningen, The Netherlands
e-mail: dick.brus@wur.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_4,

4.1 Introduction

This chapter is about statistical methods for sampling contaminated soils. I advocate statistical methods for survey of soil contamination as they provide estimated concentrations of contaminants *with quantified uncertainty*. This uncertainty can be used to optimize the number of sampling locations and/or the geographical co-ordinates of the sampling locations, i.e., the sampling pattern. From the many survey methods described in the literature, I selected some methods whose simplicity makes them attractive for application in common, real-world situations. More advanced sampling designs and estimation methods, requiring surveyors more skilled in statistics, will only be touched on.

This chapter is structured as follows. In Section 4.2 sampling methods are described for estimating the entire histogram or spatial cumulative distribution function (SCDF) of contaminant concentrations, or parameters of this SCDF such as the mean or median. Section 4.3 describes methods for estimating the means of several blocks that have been delineated before sampling. Section 4.4 describes sampling methods for mapping the concentrations of contaminants in the soil. In a final section I will go into the sampling aspects of detecting and delineating hot spots, i.e., relatively small areas where the concentration is relatively high.

The spatial resolution of the survey result increases from Section 4.2, via Sections 4.3 to 4.4. Section 4.2 considers reconnaissance surveys with the aim of estimating a single number (e.g. the mean) or several numbers (e.g. several percentiles) or the entire SCDF of the concentration for one or several contaminants. Information on where specific concentrations occur is not (yet) asked for. The other extreme is when a map is required, depicting the contaminant concentration at any point location in the study area (contaminated site) (Section 4.4). Section 4.3 is about an intermediate situation, requiring *spatial* information on the concentrations, but with less detail compared to Section 4.4: a map depicting the mean values for several blocks suffices. A typical example is estimating the means for spatial remediation units to be used for decisions on remediation of the soil in these spatial units.

This difference in spatial resolution of the survey result determines the most appropriate statistical approach for sampling. For estimating (parameters of) the spatial cumulative distribution function (SCDF), generally a *design-based* approach is most appropriate. This implies that sampling locations are selected by probability sampling (random sampling), and that in estimating (parameters of) the SCDF the selection probabilities of the sampling locations are taken into account.

For mapping purposes, a *model-based* approach is the best and only option. This implies that sampling locations need not be selected randomly, but typically are selected by purposive (targeted) sampling. In estimation (spatial interpolation) a probabilistic model of spatial variation is used.

For estimating the mean of several delineated blocks, both statistical approaches are possible. The most appropriate approach depends amongst others on the number of blocks and the affordable number of sampling locations. To estimate the means for a few blocks only, a design-based approach in which several sampling locations are randomly selected from each block may be advantageous. However, if there are many blocks, for instance more blocks than the total number of sampling locations, then a model-based approach is suggested. For more details on the choice between a design-based or a model-based sampling approach, I refer to Brus and de Gruijter (1997).

The choice between a design-based or a model-based sampling approach is only one, be it a very important decision in designing a sampling scheme. Other choices are the size and geometry (support) of the sampling units, the measurement (observation) method, the sampling design type (in the design-based approach), the spatial interpolation method (in the model-based approach), the number of sampling locations, whether or not to take composite samples *et cetera*. All these choices should be based on a thorough analysis of the aims and constraints (requirements) and on the available prior information. This is not an easy task. The development of statistically defensible environmental sampling designs that provide the right quality and quantity of data for making decisions requires a systematic planning process. This was the main motivation of the US-EPA to design a framework for this (EPA 2006). A similar approach for survey and monitoring of natural resources is described by de Gruijter et al. (2006, Part I). These frameworks can be of great help in the design process, and may prevent us from taking wrong or suboptimal decisions.

This chapter is largely based on de Gruijter et al. (2006).

4.2 Estimating (Parameters of) the Spatial Cumulative Distribution Function

This section is about how to sample if we want to estimate the spatial cumulative distribution function (SCDF) of the concentrations of one or several contaminants in the soil of an area, or parameters of this distribution function, such as the mean, median or 90th percentile. In these so-called reconnaissance surveys, it often suffices to know what concentrations occur in the area and how often, without knowing where these concentrations are. If the reconnaissance survey shows that the actual concentrations frequently exceed a specific threshold concentration, i.e., over a large proportion of the area, then we might decide to sample suspected areas additionally, so that we obtain more information on where these concentrations occur.

For estimating (parameters of) the SCDF, I generally prefer a design-based approach over a model-based approach, because in estimating the parameters of the distribution and their precision, postulation of a model of spatial variation (variogram) is not needed (Brus and de Gruijter 1997). By avoiding such a model, the quality of the estimates are independent of the quality of model-assumptions, simply because such assumptions are not needed. This property of the results is called validity. Valid results may be especially important when the survey results are used to check whether the status of the surveyed area complies with regulatory soil quality standards.

In practice, probability sampling may be hampered by enclosures that are part of the target population but are inaccessible for sampling, e.g. under sealed parts of the area. If a randomly selected location happens to fall in such an enclosure, or if in the field it appears that the location falls outside the target population, this location must be eliminated from the list of sampling locations, and replaced by a new sampling location from a reserve list in pre-determined order. Shifting such locations to nearby accessible locations inside the target population may cause biased estimates.

Hereafter, I will describe several sampling designs (Section 4.2.1), and how (parameters of) the SCDF can be estimated from the sample (Section 4.2.2). In Section 4.2.3. I will describe how ancillary information such as maps with covariates, or existing measurements of the contaminant concentrations at point locations can be used at the estimation stage of a sampling strategy. Section 4.2.4 describes composite sampling, in which the soil taken at different sampling locations is mixed (bulked) to save laboratory costs. The final Section 4.2.5 deals with the decision on the number of sampling locations.

4.2.1 Sampling Designs

There are many sampling designs described in textbooks that are appropriate for surveying contaminated soils, see for instance de Gruijter et al. (2006). Here, I will describe only a few simple sampling designs that are applicable in many common situations. Section 4.2.1.4 touches on other, more advanced sampling designs, and describes when these alternative designs can be advantageous. I prefer simple sampling designs to keep estimation as simple as possible. Also, if one has many soil contaminants to be surveyed, we may optimize the sampling design for one or two contaminants, but this design can be far from optimal for other contaminants. For instance, the precision of the estimated mean concentration of a contaminant may be considerably increased by selecting sampling locations with probabilities proportional to a prior estimate of the concentration. However, for other contaminants poorly correlated with the first contaminant, this so-called pps-sampling design can be rather inefficient, see Brus et al. (2006) for an example. Finally, ancillary information either can be used at the sampling stage, or at the estimation stage, see Section 4.2.3. The latter is more flexible: we can use this information to estimate the mean concentration of some contaminants but not for others.

Many of the sampling designs described hereafter are supported by Visual Sample Plan (VSP) software for designing environmental sampling plans. VSP can be downloaded free at http://vsp.pnl.gov, along with a user's guide (Matzke et al. 2007).

4.2.1.1 Simple Random Sampling

The simplest way of selecting sampling locations randomly is simple random sampling (SI). In SI all sampling locations are selected with equal probability and independently from each other. In general, the sampling locations are slightly clustered by chance. This makes SI rather inefficient in general: the sampling variance (standard deviation) of the estimated target parameter (mean, median, areal proportion et cetera.) is large compared to other sampling designs with the same number of sampling locations. An advantage of SI is that it is easy to implement, see de Gruijter et al. (2006, p. 80), and the estimation of the target parameter and its sampling variance is relatively simple (see Section 4.2.2).

4.2.1.2 Stratified Simple Random Sampling

In stratified simple random sampling (STSI) the area is divided into sub-areas, called 'strata', in each of which a predetermined number of sampling locations is selected by SI, see previous section. There are two possible reasons for stratification. First, by stratifying the area we may aim at increasing the efficiency compared to SI. So, due to the stratification we hope that the sampling variance of the estimated target parameter is smaller than with simple random sampling at the same costs (same number of sampling locations), or vice versa, we hope that the costs of STSI are smaller (we need fewer sampling locations) compared to SI with the same sampling variance. The more homogeneous the strata, the larger the gain in efficiency. A homogeneous stratum is a stratum in which the soil contaminant concentration varies only slightly compared to the spatial variation within the area as a whole. A pitfall in stratified sampling is to use too many strata. This is no problem as long as the number of sampling locations per stratum is nearly proportional to the surface area. However, if this detailed stratification leads to numbers of sampling locations per stratum that are strongly disproportional to the surface areas, this may lead to a loss rather than a gain in precision.

The second reason for stratification is that we want to have separate estimates for the sub-areas, see Section 4.3. By using the sub-areas as strata in random sampling, we can control the number of sampling locations within the sub-areas, and related to this control the precision of the estimated parameters of the SCDF. Note that by not using the sub-areas as strata in sampling, the number of sampling locations within a given sub-area is not fixed, but random, i.e. it varies between samples drawn with the same sampling design.

Figure 4.1 illustrates a STSI sample of 100 locations for estimating the SCDF of 252 chemical compounds in the topsoil (0–10 cm) and subsoil (50–100 cm)

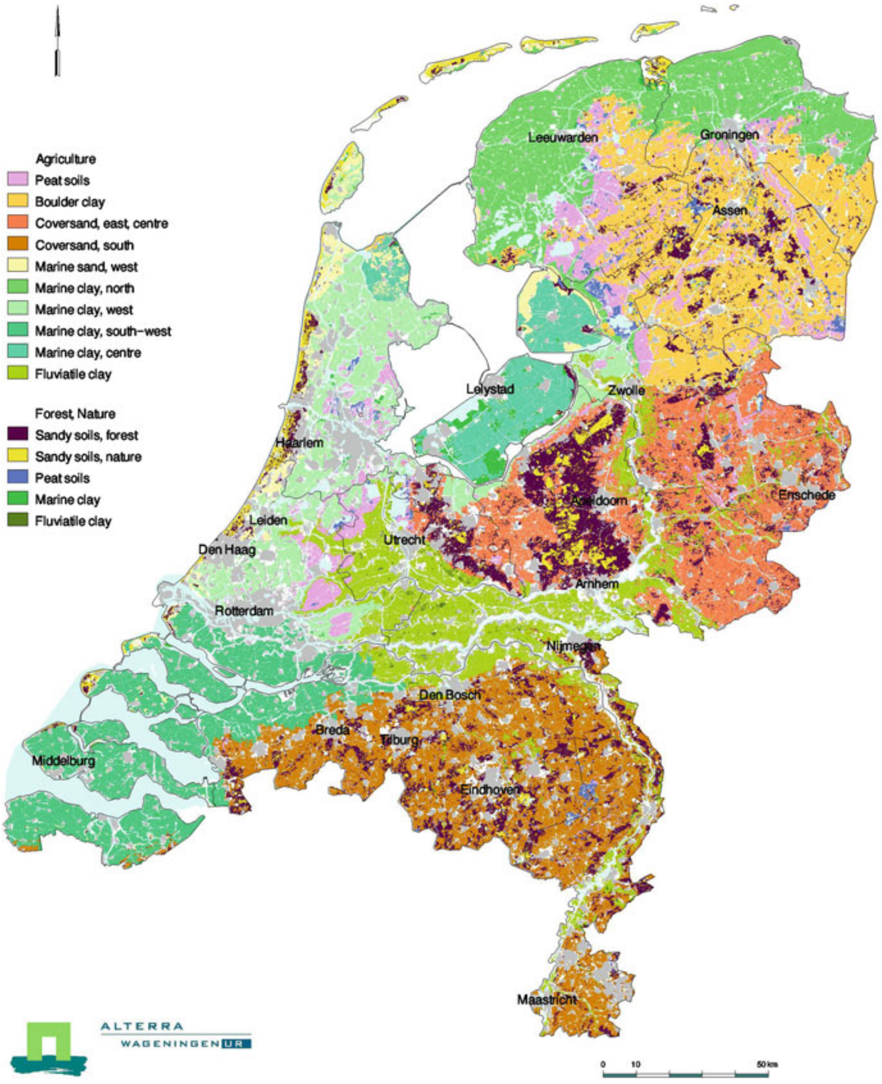


Fig. 4.1 (continued)

Fig. 4.1 (continued) Strata used to estimate the SCDF of soil contaminants in the Netherlands (*left*), and selected sample locations (*right)* (details not clearly visible; both graphs are meant to give a general insight into soil maps)

of the Netherlands (Brus et al. 2009). To increase the efficiency, the study area is stratified according to soil type and land use. The two largest strata, “cultivated sandy soils” and “cultivated marine clay soils” cover large parts of the Netherlands. Therefore, these two strata were subdivided into four geographical substrata each, which improved the spatial coverage of the study area. Finally, the stratum “uncultivated sandy soils” is split into “sandy forest soils”, and “other uncultivated sandy soils” (dunes, heath). The total number of strata thus obtained was 15. The number of sampling locations within a stratum is approximately proportional to its area. The minimum number of sampling locations was two. This is to avoid problems with estimating the sampling variance, see Section 4.2.2.

Geographical Stratification

If we have no prior information on the spatial distribution of the soil contaminant that can be used for stratification, then we may use geographical strata to spread the sampling locations as uniformly as possible over the study area. Spreading the sampling locations over the area generally enhances the precision of the estimated parameters. The best spatial coverage will be obtained when the geographical strata have maximum compactness. A simple, straightforward method for computing such compact geographical strata is k-means clustering of small grid cells, using the x- and y-coordinates of their centres as classification variables, see Brus et al. (1999) for more details. Compact geographical strata can be designed by the R-package spcosa (Walvoort et al. 2009, 2010).

Figure 4.2 shows an area of 836 ha with peat soils near Mijdrecht, south of Amsterdam in the Netherlands, planned for development of nature. Since the reclamation of the peat areas in the Middle Ages, a mixture of garbage from the nearby cities, farmyard manure, dune-sand and dredged sediments, has been dumped on the weak peat soils. As a result, the topsoils are contaminated with Pb, Cu and Zn. The study area was divided into 15 compact geographical strata of equal surface area. In each stratum, two sampling locations were selected by simple random sampling.

4.2.1.3 Random Grid Sampling

A simple way of drawing random samples with good spatial coverage, i.e. samples whose locations are spread uniformly over the study area, is random grid sampling (SY). Possible shapes for the grid cells are square, triangular, and hexagonal. The triangular shape was shown to be most efficient in general. Besides the shape of the grid cells, we must decide on their size (grid spacing) and on the orientation of the grid. The grid spacing determines the number of sampling locations in the study area. So, if we have decided on the required (allowed) number of sampling locations, then we may use this number to calculate the grid spacing. For square grids, the grid spacing (m) can be calculated with $\sqrt{A/n}$, where A is the area in m^2, and n is the number of sampling locations. The grid is randomly placed over the study area as follows. One location is selected by simple random sampling from the study area. Given the chosen orientation of the grid, the grid is extended in all directions using the selected location as a starting node. Finally, all nodes are selected that fall within the study area. There is no need for random selection of the orientation of the grid, random selection of the first node suffices for design-based statistical inference (estimation).

In general, the spatial coverage of an area by a random grid is better than by a geographically stratified random sample, even with one location per stratum. Consequently, in general, random grid sampling will give more precise estimates of the (parameters of) the SCDF. There are two disadvantages of random grid sampling compared to geographically stratified random samples. First, estimation of the sampling variance is cumbersome. This is because we do not have independent replicates of the sample: the grid can be considered as one ‘cluster’ of sampling locations, see Section 4.2.2.1 Second, in general the number of sampling locations

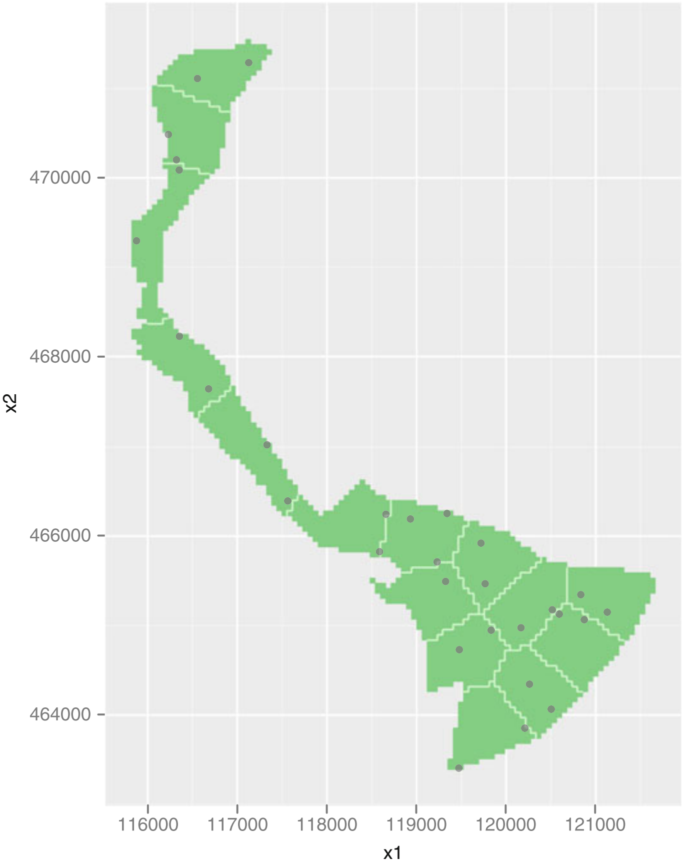


Fig. 4.2 Stratified simple random sample of 30 locations in peat soils near Mijdrecht, south of Amsterdam, polluted by Pb, Cu and Zn. The strata are compact blocks of equal surface area, formed by k-means clustering of the cells of a fine grid. The geographical stratification enhances the spatial coverage of the study area by the sample. Compact geographical strata were computed by R-package spcosa (Walvoort et al. 2009, 2010)

with random grid sampling is not fixed, but varies between randomly drawn samples. We may choose the grid spacing such that, on average, the number of sampling locations equals the required (allowed) number of sampling locations, but for the actually drawn sample, this number can be a few locations smaller or larger. A random number of sampling locations may be undesirable, for instance, when this size is prescribed in regulations.

4.2.1.4 Advanced Sampling Designs

There are various, more advanced sampling designs described in handbooks on sampling theory that have potentials for survey of soil contamination. I will describe two of these sampling designs, double sampling and adaptive cluster sampling.

Double Sampling

In double sampling, also referred to as two-phase sampling or collaborative sampling, two methods for measuring the contaminant concentrations are combined, i.e., a cheap and relatively imprecise method, and an expensive, relatively precise method. In many cases the cheap measurements are collected in the field, whereas the expensive ones are obtained by laboratory analysis. In double sampling, a large probability sample is selected where the contaminant concentration is measured with the cheap method. From this large "first-phase" sample a *random* subsample is selected where the contaminant concentration is also measured with the expensive method. If the concentrations measured with both methods are strongly correlated, then the efficiency of the sampling method can be increased. The optimal ratio of the two numbers of sampling locations is

$$\frac{m}{n} = \sqrt{\frac{c_z}{c_y}\frac{1-\rho^2}{\rho^2}}, \tag{4.1}$$

with m and n the numbers of sampling locations with expensive and cheap measurements, respectively, and c_y and c_z the measurement costs per location for the expensive and cheap method, respectively. Chang and Yeh (2007) illustrate this sampling method for estimating mean heavy metal concentrations, using relatively cheap, field portable x-ray fluorescence measurements besides standard laboratory measurements. Double sampling is cost-effective compared to measuring with the precise method only if (Cochran 1977, p. 341)

$$\frac{c_y}{c_z} > \frac{\left(1+\sqrt{1-\rho^2}\right)^2}{\rho^2}. \tag{4.2}$$

Cheap measurements of the contaminant concentration can also be replaced by measurements on a continuous covariate correlated with the contaminant concentration, such as hand estimates of the clay percentage or the organic matter content. This covariate can also be a categorical variable, for instance classes of contamination level (e.g. not contaminated, moderately contaminated, strongly contaminated) based on organoleptic observations.

The mean concentration can be estimated by the regression estimator (continuous covariate) or post-stratification estimator (categorical covariate), see Section 4.2.3 In this case, the known spatial mean for the covariate in Eq. (4.24) must be replaced by the mean estimated from the large "first-phase" sample, and the relative area in Eq. 4.22 by the ratio of the number of sampling locations in a given category and the total number of locations of the "first-phase" sample, n.

Adaptive Cluster Sampling

Adaptive cluster sampling is a type of sequential sampling where additional sampling locations are selected in the neighbourhood of sampling locations that

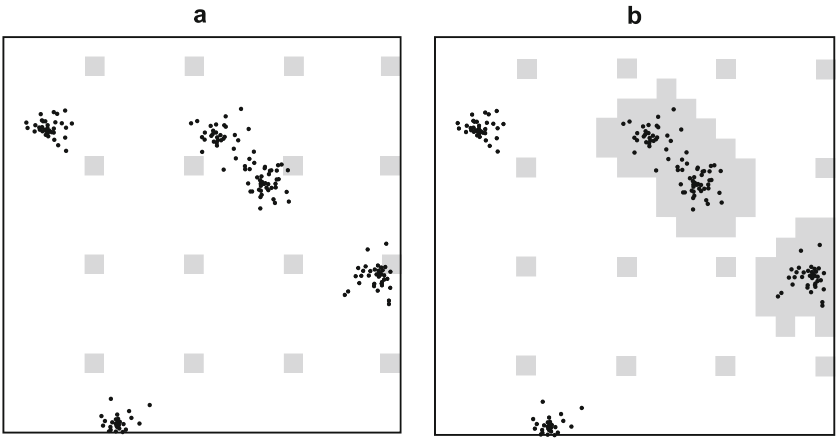


Fig. 4.3 Adaptive cluster sample. The initial sample is a random grid sample of 16 square plots (**a**). In the successive stages the four neighbouring plots of a selected plot are added to the sample if this plot contains at least one polluted site (**b**)

satisfy a specific criterion, for instance concentrations exceeding a specific threshold concentration, e.g. a soil quality standard (Fig. 4.3). This type of sequential sampling leads to spatial clustering of sampling locations. This can be efficient if locations with concentrations exceeding the threshold concentration are strongly clustered. For details I refer to Thompson and Seber (1996).

4.2.2 Estimation

The formulas for estimating the (parameters of) the SCDF and for estimating their sampling variances generally differ between the sampling designs. I have seen several studies in which some "smart" sampling design was applied, but after collection of the data, the surveyors apparently forgot about this design, and estimated the parameters and their variances as if it were a SI sample. This is not proper science. The sampling design determines the selection (inclusion) probabilities of sampling locations and of pairs of sampling locations, and these selection probabilities must be accounted for in estimating the parameters and their variances. Hereafter, formulas for estimating the spatial mean (Section 4.2.2.1), areal fraction (proportion of the area) with concentrations exceeding a threshold concentration (Section 4.2.2.2), SCDF (Section 4.2.2.3), and percentiles (Section 4.2.2.4) are given for the three sampling designs described in the previous section.

4.2.2.1 Spatial Mean

Simple Random Sampling

For SI the spatial mean is estimated by the unweighted mean of the measurements in the sample

$$\hat{\bar{y}}_{\mathrm{SI}} = \frac{1}{n}\sum_{i=1}^{n} y_i \;, \tag{4.3}$$

where n is the number of selected sampling locations, and y_i is the contaminant concentration at sampling location i. The subscript SI is added to stress that this estimator is intended for simple random sampling designs.

The *sampling* variance of the estimated mean is estimated by

$$\widehat{V}\left(\hat{\bar{y}}_{\mathrm{SI}}\right) = \frac{\widehat{S^2}(y)}{n} \;, \tag{4.4}$$

where $\widehat{S^2}(y)$ is the estimated *spatial* variance of the concentrations in the study area

$$\widehat{S^2}(y) = \frac{1}{(n-1)}\sum_{i=1}^{n}\left(y_i - \hat{\bar{y}}_{\mathrm{SI}}\right)^2 \;. \tag{4.5}$$

The $100(1-\alpha)\%$ confidence interval for $\bar{y}$ is given by:

$$\hat{\bar{y}}_{\mathrm{SI}} \pm t_{1-\alpha/2}^{(n-1)} \cdot \sqrt{\widehat{V}\left(\hat{\bar{y}}_{\mathrm{SI}}\right)} \;, \tag{4.6}$$

where $t_{1-\alpha/2}^{(n-1)}$ is the $(1-\alpha/2)$ quantile of the Student distribution with $(n-1)$ degrees of freedom. This confidence interval is based on the assumption that y, and as a consequence $\hat{\bar{y}}_{\mathrm{SI}}$, is normally distributed. If the distribution deviates clearly from normality, the data should be first transformed to normality, for instance by taking the logarithm. The confidence limits thus found are then back-transformed to the original scale. However, we must be aware that the confidence intervals thus obtained are not the confidence intervals of the mean on the original scale. For instance, for a lognormal distribution, the confidence interval is for the median on the original scale. A simple rule to obtain the confidence limits of the estimated mean is to multiply the back-transformed limits by the ratio of the untransformed mean and the back-transformed mean (R.L. Correll, personal communication). Transformation is not necessary if n is large, because $\hat{\bar{y}}_{\mathrm{SI}}$ is then approximately normally distributed according to the Central Limit Theorem.

Stratified Simple Random Sampling

For STSI the spatial mean is estimated by

$$\hat{\bar{y}}_{\mathrm{STSI}} = \sum_{h=1}^{L} a_h \, \hat{\bar{y}}_h \;, \tag{4.7}$$

where L is the number of strata, a_h are the relative areas of the strata (stratum weights), and $\hat{\bar{y}}_h$ is the estimated mean of stratum h estimated by Eq. (4.3).

Provided all numbers of sampling locations are larger than 1, the sampling variance of $\hat{\bar{y}}_{\mathrm{STSI}}$ can be estimated by

$$\widehat{V}\left(\hat{\bar{y}}_{\mathrm{STSI}}\right) = \sum_{h=1}^{L} a_h^2 \widehat{V}\left(\hat{\bar{y}}_h\right) , \tag{4.8}$$

where $\widehat{V}\left(\hat{\bar{y}}_h\right)$ is the estimated sampling variance of $\hat{\bar{y}}_h$

$$\widehat{V}\left(\hat{\bar{y}}_h\right) = \frac{S_h^2(y)}{n_h} , \tag{4.9}$$

with $S_h^2(y)$ the *spatial* variance of y within stratum h, that can be estimated by (compare with Eq. (4.5))

$$\widehat{S^2}_h(y) = \frac{1}{(n_h - 1)} \sum_{i=1}^{n_h} \left(y_{hi} - \hat{\bar{y}}_h\right)^2 , \tag{4.10}$$

and n_h is the number of sampling locations in stratum h.

If only one location per stratum is selected, the sampling variance can be approximated by the collapsed strata estimator (Cochran 1977).

The $100(1-\alpha)\%$ confidence interval for $\bar{y}$ is given by

$$\hat{\bar{y}}_{\mathrm{STSI}} \pm t_{1-\alpha/2}^{(df)} \cdot \sqrt{\widehat{V}\left(\hat{\bar{y}}_{\mathrm{STSI}}\right)} , \tag{4.11}$$

where $t_{1-\alpha/2}^{(df)}$ is the $(1-\alpha/2)$ quantile of the Student distribution with *df* degrees of freedom. For *df* we can either take simply $n{-}L$ (Lohr 1999, p. 101), or it can be approximated by Sattherwaites method (Nanthakumar and Selvavel 2004)

$$df \approx \frac{\left(\sum_{h=1}^{L} a_h^2 \frac{\widehat{S^2}_h(y)}{n_h}\right)^2}{\sum_{h=1}^{L} a_h^4 \left(\frac{\widehat{S^2}_h(y)}{n_h}\right)^2 \frac{1}{n_h - 1}} \tag{4.12}$$

Random Grid Sampling

Similar to SI, for SY the spatial mean is estimated by the unweighted mean of the measurements in the sample

$$\hat{\bar{y}}_{\mathrm{SY}} = \frac{1}{n} \sum_{i=1}^{n} y_i . \tag{4.13}$$

As mentioned above, estimation of the sampling variance of $\hat{\bar{y}}_{\mathrm{SY}}$ is cumbersome. A simple, often applied procedure is to calculate the sampling variance as if the sample is a simple random sample, i.e. the variance is estimated by Eq. (4.4). In general this procedure over-estimates the sampling variance, so that we are on the safe side. Alternatively, if a variogram can be estimated from the random grid sample, then this variogram can be used to estimate the sampling variance, see de Gruijter et al. (2006, p. 131) for details.

For SY, confidence intervals can be approximated by Eq. (4.6).

4.2.2.2 Areal Fraction with Concentrations Exceeding Threshold Concentration

The above formulas for estimating means can also be used to estimate the *fraction* (proportion) of the area with concentrations exceeding a given threshold concentration. The measured concentrations at the sampling locations $j = 1 \cdots n$ are transformed to a 0/1 indicator variable, with value 1 if the concentration exceeds the threshold and 0 otherwise

$$i_{j,y_\mathrm{t}} = \begin{cases} 1 & \text{if} \quad y_j > y_\mathrm{t} \\ 0 & \text{if} \quad y_j \leq y_\mathrm{t} \end{cases} \tag{4.14}$$

where y_t is the threshold concentration. The above formulas are then simply applied to this indicator variable. For SI and SY this boils down to computing the proportion of sampling locations with concentrations exceeding the threshold

$$\hat{p}_{\mathrm{SI}} = \hat{p}_{\mathrm{SY}} = \frac{1}{n} \sum_{j=1}^{n} i_{j,y_\mathrm{t}} \, . \tag{4.15}$$

For STSI the areal fraction is estimated as a weighted mean of the sample proportions per stratum, using the relative areas of the strata as weights.

For SI the sampling variance of the estimated areal fraction is estimated by

$$\widehat{V}\big(\hat{p}_{\mathrm{SI}}\big) = \frac{\widehat{S^2}(i_{y_\mathrm{t}})}{n-1} \, , \tag{4.16}$$

where $\widehat{S^2}(i_{y_\mathrm{t}})$ is the estimated spatial variance of the indicator

$$\widehat{S^2}(i_{y_\mathrm{t}}) = \hat{p}_{\mathrm{SI}}(1 - \hat{p}_{\mathrm{SI}}). \tag{4.17}$$

For STSI the sampling variance of the estimated fraction is estimated by

$$\widehat{V}(\hat{p}_{\mathrm{STSI}}) = \sum_{h=1}^{L} a_h^2 \frac{\widehat{S^2}_h(i_{y_\mathrm{t}})}{n_h - 1} \, , \tag{4.18}$$

where $\widehat{S^2}_h(i_{y_\mathrm{t}})$ is the estimated spatial variance of the indicator within a stratum

$$\widehat{S^2}_h(i_{y_t}) = \hat{p}_h(1 - \hat{p}_h). \tag{4.19}$$

For large numbers of sampling locations, confidence intervals for areal fractions can be estimated by assuming a normal or Student distribution, see Eqs. (4.6) and (4.11). Problems arise for small samples (say $n < 20$), and when estimating extreme (very small or very large) areal fractions. In these cases exact confidence bounds based on the binomial distribution are required. For SI such exact bounds can be easily calculated. However, for STSI this is not straightforward (Wendell and Schmee 1996). For STSI confidence bounds can be approximated by assuming the standardized estimated areal fraction has a Student distribution

$$\frac{\hat{p}_{\mathrm{STSI}} - p}{\sqrt{\widehat{V}(\hat{p}_{\mathrm{STSI}})}} \sim t^{(df)}, \tag{4.20}$$

with the degrees of freedom *df* equal to Eq. (4.12). The $100(1-\alpha/2)\%$ interval for the areal fraction is

$$\hat{p}_{\mathrm{STSI}} \pm t^{(df)}_{1-\alpha/2}\sqrt{\widehat{V}(\hat{p}_{\mathrm{STSI}})}. \tag{4.21}$$

For SY, confidence bounds can be approximated as for STSI.

4.2.2.3 Spatial Cumulative Distribution Function

The SCDF of y can be estimated through repeated application of the indicator technique described in Section 4.2.2.2. The measured variable y is first transformed to a *series* of indicator variables corresponding to a number of increasing threshold concentrations. Note that, contrary to Eq. (4.14) in this case the indicator has a value of 1 if the concentration is *smaller than or equal to* the threshold, and a value of 0 else. The areal fractions estimated from these indicators, together with the threshold concentrations, form an estimate of the SCFD. It is common to use all different concentrations in the sample as thresholds. If there are measurements below a detection limit, then this detection limit is the smallest threshold. If there are several detection limits, the largest detection limit is used as the first threshold.

Figure 4.4 shows the SCDF's of the Cu, Pb, Co and Ni concentrations in the topsoils (0–10 cm, brown diamonds) and subsoils (50–100 cm, yellow squares) in the Netherlands, estimated from the stratified simple random sample depicted in Fig. 4.4. The Cu and Pb concentrations in topsoils are higher than in subsoils due to anthropogenic contamination, whereas Co and Ni have geogenic origin, and consequently concentrations in topsoils and subsoils are comparable.

Figure 4.5 shows the estimated SCFD of Pb for the peat soils near Mijdrecht. Six of the 30 sampling locations from three strata (Fig. 4.2) could not be sampled (no permission). The Pb concentrations are much higher than in the Netherlands as a whole, compare with Fig. 4.4.

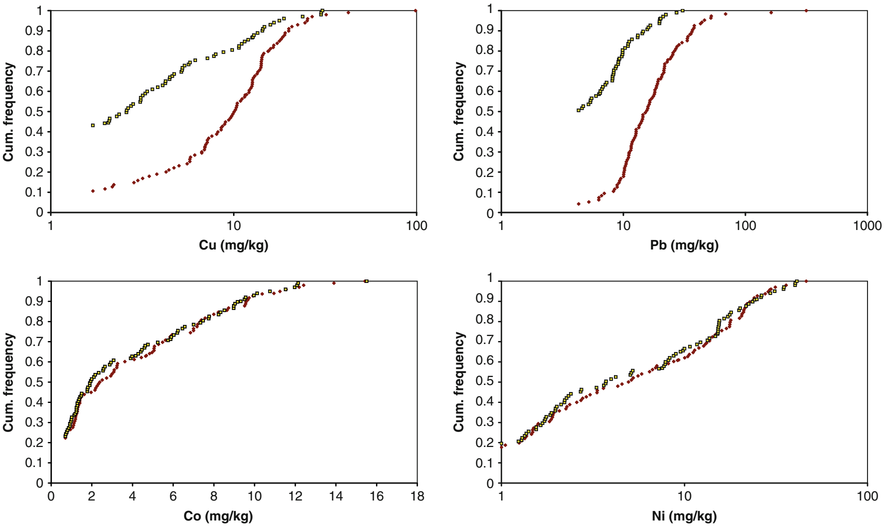


Fig. 4.4 Estimated spatial cumulative distribution functions of Cu, Pb, Co and Ni in topsoils of the Netherlands

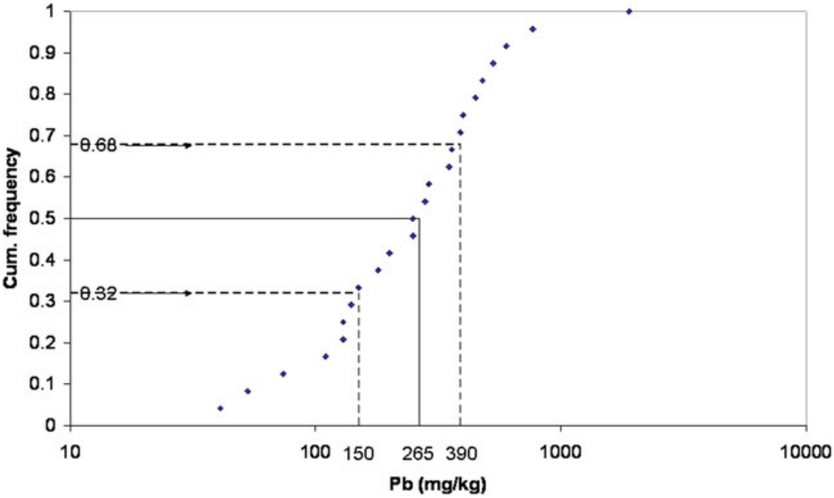


Fig. 4.5 Estimation of median Pb concentration and its 95% confidence interval in peat soils at Mijdrecht from stratified simple random sample, see Fig. 4.2

4.2.2.4 Median and Other Percentiles

The median and any other percentile (quantile) can be estimated by inverse use of the estimated SCDF. Confidence bounds for the estimated percentiles can be obtained via the estimated confidence bounds for the cumulative frequencies, see

Fig. 4.5 (Särndal et al. 1992). The 95% confidence interval for the cumulative frequency has been computed with the Student distribution, approximating the degrees of freedom by Eq. (4.12).

4.2.3 Using Ancillary Information in Estimation

Ancillary information such as maps with covariates either can be used at the sampling stage or at the estimation stage of a sampling strategy. An example of use at the sampling stage is stratified random sampling. This section describes how ancillary information can be used at the estimation stage. The formulas (estimators) in Section 4.2.2 do not make use of such ancillary information.

4.2.3.1 Post-Stratification Estimator

If the study area can be split up into sub-areas that have less variation within them than in the area as a whole, the efficiency can be increased by using these sub-areas either at the sampling stage (stratified sampling) or at the estimation stage. In the latter case, for each sampling location the sub-area (in sampling terminology referred to as a group) must be determined. For SI, the mean (areal fraction) can then be estimated by the post-stratification estimator

$$\hat{\bar{y}}_{\mathrm{pos}} = \sum_{g=1}^{G} a_g \, \bar{y}_{\mathrm{s}_g} \,, \tag{4.22}$$

where a_g is the relative area of sub-area g, and $\bar{y}_{\mathrm{s}_g}$ is the sample mean of sub-area g.

In the case of a stratified simple random sample, for which one wants to use a second grouping at the estimation stage, the mean can be estimated by

$$\hat{\bar{y}}_{\mathrm{pos}} = \sum_{g=1}^{G} a_g \, \hat{\bar{y}}_g = \sum_{g=1}^{G} a_g \sum_{h=1}^{L_g} \frac{\widehat{A}_{gh}}{\widehat{A}_g} \, \bar{y}_{\mathrm{s}_{gh}} \,, \tag{4.23}$$

where L_g is the number of strata in group g, $\widehat{A}_g$ is the estimated area of group g, and $\widehat{A}_{gh}$ and $\bar{y}_{\mathrm{s}_{gh}}$ are the estimated surface area and the sample mean of group g in stratum h, respectively. Note that the relative sizes a_g must be known. Also note that Eq. (4.23) uses the ratio of the estimated areas $\widehat{A}_{gh}$ and $\widehat{A}_g$. This is because this gives more precise estimates than the ratio of the true areas.

4.2.3.2 Regression-Estimator

If *quantitative* ancillary information is available and is known everywhere in the study-area, for instance from remote sensing or from a digital terrain model, then the spatial mean (areal fraction) can be estimated by the regression estimator

$$\hat{\bar{y}}_{\text{reg}} = \hat{\bar{y}}_{\pi} + b\left(\bar{z} - \hat{\bar{z}}_{\pi}\right) , \tag{4.24}$$

where

$\hat{\bar{y}}_{\pi}$ is the spatial mean of the soil contaminant y estimated from the measurements of y in the probability sample only;
$\hat{\bar{z}}_{\pi}$ is the estimated mean of the ancillary variable;
$\bar{z}$ is the true mean of the ancillary variable; and
b is the estimated regression coefficient (slope) for the ancillary variable.

The estimators $\hat{\bar{y}}_{\pi}$ and $\hat{\bar{z}}_{\pi}$ are the design-specific estimators for the mean, presented in Section 4.2.2.

As with spatial means, in estimating the slope the sampling design must be taken into account. For SI, the slope can be estimated by the least squares estimator

$$b = \frac{\sum_{i=1}^{n}(y_i - \hat{\bar{y}}_{\text{SI}})(z_i - \hat{\bar{z}}_{\text{SI}})}{\sum_{i=1}^{n}(y_i - \hat{\bar{y}}_{\text{SI}})^2} . \tag{4.25}$$

For STSI, Eq. (4.25) is used to estimate the slopes per stratum. If the number of sampling locations selected in a stratum is small, e.g. $n_h < 10$ to 20 (depending on the number of regression coefficients), this stratum must be combined with others to obtain valid estimates of the sampling variance. For a combination of strata, the regression coefficients can be estimated by the weighted means of the coefficients per stratum, using the relative areas as weights (Eq. 4.7). For SY, the slope can be simply estimated by Eq. (4.25).

An interesting application of the regression estimator is the use of legacy data *not selected by probability sampling* in unbiased estimation of spatial means (Brus and de Gruijter 2003). For instance, one may have legacy data on a soil contaminant, possibly preferentially sampled at contaminated or uncontaminated sites. In the method proposed by Brus and de Gruijter (2003) a relatively small probability sample is added to the non-probability sample. The concentrations measured at the legacy sampling locations are interpolated to these new sampling locations. The interpolated variable is then used as a covariate in the regression estimator.

4.2.4 Composite Sampling

Laboratory costs can be saved by taking composite samples, i.e. by bulking the soil aliquots taken at individual sampling locations, and mixing them thoroughly (Boswell et al. 1996; Elder et al. 1980; Rohde 1976). To profit fully from compositing it is necessary to mix and homogenize the subsamples (increments) that are pooled together to form a composite sample, to the extent that the spatial variation between the increments is eliminated. In practice this can be hard to achieve, so that it can become efficient to repeat the subsampling of the composite for laboratory analysis.

This technique is useful if the aim is to estimate the spatial mean, but for other aims such as estimating a percentile, the entire SCFD, or the areal fraction exceeding a threshold concentration it can not be used. The variation of the concentration in the composite sample over repeated sampling will be small, because the spatial variation between the aliquots of a composite sample is eliminated. The larger the spatial variation between aliquots, the more precise the composite sample mean. Bulking soil aliquots from different strata therefore pays. An example is where the study area is geographically stratified (see Section 4.2.1.2), and one aliquot is taken from each stratum to form a composite sample. Note that equal volumes of soil must be collected at each location and the strata must have an equal surface area in order to obtain unbiased estimates. If an estimate of the sampling variance of the estimated mean is required, then several composite samples must be taken. If multiple composite samples are taken, the spatial mean is estimated by the average of the composite sample means

$$\widehat{\bar{y}}_{\mathrm{com}} = \frac{1}{C}\sum_{i=1}^{C} y_i; , \tag{4.26}$$

where C is the number of composite samples, and y_i the (average) concentration of composite sample i. The sampling variance of this estimated mean equals

$$\widehat{V}(\bar{y}_{\mathrm{com}}) = \frac{\widehat{V}(y_i)}{C}. \tag{4.27}$$

Another application of composite sampling is group screening or group testing. In group screening the aim is to determine whether, for instance, a contaminant or a species of soil microbe, is present or not. The detection limit of the method for (chemical) analysis must be low enough to detect the contaminant or species in a (strongly) diluted sample.

4.2.5 *Required Number of Sampling Locations*

As stated in the Introduction (Section 4.1) the advantage of statistical methods for survey is that the precision of the survey results (estimates) can be quantified. In the formulas for the sampling variance, see Eqs. (4.4), (4.8), (4.15) and (4.18), *spatial* variances appear, that are estimated from the same sample. If we are able, prior to the sampling, to make a first guess at these spatial variances, we can compute in advance the number of sampling locations required to achieve a given minimum quality of the estimated target parameter. The computations described hereafter are supported by the software Visual Sample Plan (VSP) (Matzke et al. 2007).

4.2.5.1 Constraint on Sampling Variance or Coefficient of Variation

The simplest situation is when the quality constraint is formulated in terms of the sampling variance or standard deviation. For SI, the required number of sampling locations can then simply be obtained by rewriting Eqs. (4.4) and (4.16), and substituting the prior estimate of the spatial variance of y or the indicator i in the area, $\breve{S}^2(.)$, for the estimated spatial variance, $\widehat{S^2}(.)$

$$n = \frac{\breve{S}^2(.)}{V_{\max}} \;. \tag{4.28}$$

In case of a constraint on the coefficient of variation, the same procedure can be used. It requires a prior estimate of the mean, which is used to compute the maximum value of the sampling variance of the mean, $V_{\max}$.

For STSI, prior estimates of the spatial variances within strata can be used to compute the optimal allocation, i.e. the numbers of sampling locations per stratum that result in the minimum sampling variance of the estimated mean or areal fraction, given the total number of sampling locations (de Gruijter et al. 2006, p. 94). The required total number of sampling locations for a specified maximum allowable sampling variance of the estimated mean (areal fraction) can then be calculated with

$$n = \frac{1}{V_{\max}} \cdot \left(\sum_{h=1}^{L} a_h \, \breve{S}_h(.) \right)^2 \;, \tag{4.29}$$

where $\breve{S}_h(.)$ is the prior estimate of the spatial standard deviation of y or i within stratum h. Equation (4.29) assumes equal cost per sampling location for the strata. I refer to de Gruijter et al. (2006, p. 95) if these costs are unequal for the strata. For proportional allocation, the required number of sampling locations can be calculated with

$$n = \frac{\sum\limits_{h=1}^{L} a_h \, \breve{S}_h^2(.)}{V_{\max}} \;. \tag{4.30}$$

4.2.5.2 Constraint on Probability of Error

This section describes how to calculate the required number of sampling locations if the quality constraint is formulated in terms of the probability of occurrence of the error in the estimated mean or the areal fraction. If an absolute error d has been specified with an allowed probability of exceedance α, then the maximum allowable sampling variance can be derived from d and α as

$$V_{\max} = \left(\frac{d}{u_{1-\alpha/2}} \right)^2 \;, \tag{4.31}$$

where $u_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution. The derived value of $V_{\max}$ is then inserted in Eq. (4.28), Eq. (4.29) or Eq. (4.30) to obtain the required number of sampling locations.

If instead of the absolute error, the relative error $|(\hat{\bar{y}} - \bar{y})/\bar{y}|$ should be smaller than a specified limit r, the required number of sampling locations can be used by replacing d in the numerator of Eq. (4.31) by $r\,\breve{\bar{y}}$, where $\breve{\bar{y}}$ is a prior estimate of the mean, and continuing the procedure as before.

For small numbers of sampling locations ($n < 20$), we must use the Student distribution, see Eqs. (4.6) and (4.11), to compute $V_{\max}$. Now we face the problem that we must know the number of sampling locations for SI, and the numbers of sampling locations per stratum for STSI (see Eq. (4.12)) in order to compute the number of degrees of freedom *df*, and subsequently $t_{1-\alpha/2}^{(df)}$. In this situation we recommend calculating, for a series of total number of sampling locations, the cumulative (lower) probability of $\frac{d}{\sqrt{\breve{V}(y)/n}}$ for the Student distribution with *df* degrees of freedom. In this case, the required number of sampling locations is the smallest number for which the cumulative probability is smaller than $\alpha/2$.

The number of sampling locations needed to estimate an areal fraction p such that, with a specified large probability $(1 - \alpha)$, the absolute or relative error in the estimated fraction is smaller than a particular limit, can be calculated with the binomial distribution. This requires a prior estimate of the areal fraction only. For STSI, use of the binomial distribution for computing the required number of sampling locations is much more complicated, and I recommend using the Student distribution for this purpose, see Eq. (4.21).

4.2.5.3 Constraints on Error Rates in Testing of a Hypothesis

If the survey results are used in decision-making by means of statistical testing of a hypothesis on the spatial mean or the areal fraction, then the quality constraint is best formulated in terms of the probability of false rejection (type I error) and the probability of false acceptance (type II error) of the hypothesis. An example is testing the spatial mean concentration against a soil quality standard (compliance monitoring). Both probabilities (error rates) are set to a maximum. The complement of the maximum probability of false acceptance, $1 - \beta$ with β the probability of type II error, is referred to as the power of the test. This power must be linked to a "minimum detectable difference". Note that a hypothesis on a percentile can be reformulated as a hypothesis on the areal fraction. For instance, testing the null-hypothesis "The 95th-percentile of Zn $\leq$ 140 mg kg^{-1}" is equivalent to testing the null-hypothesis "The areal fraction with Zn concentrations $\leq$ 140 mg kg^{-1} $\geq$ 0.95". I refer to de Gruijter et al. (2006, pp. 85–89) and to EPA (2006) for more details on this subject matter.

4.3 Estimating Mean Concentrations for Delineated Blocks

This section describes sampling strategies for estimating the mean contaminant concentrations of several blocks that are delineated before sampling. The decision on remediation of a block is typically based on estimates of the block-mean contaminant concentration. Contrary to the previous section, an estimate of the mean for the study area as a whole (global mean) or of its SCDF, is not sufficient for this purpose. For remediation purposes we need to know where the concentrations exceed threshold concentrations. As remediation measures are generally applied on blocks with dimensions in the horizontal plane of tens to hundreds of m^2, the required spatial resolution of the map can be adapted to these dimensions, and estimation at point locations (see Section 4.4) is not needed.

To estimate the mean concentrations for the delineated blocks, either a design-based or a model-based approach can be applied. The best sampling strategy depends, amongst others, on the number of blocks and the affordable number of sampling locations or, in case composite samples are taken, the affordable number of composite samples. If we can afford several sampling locations (composite samples) in each block, say more than five, then a design-based approach can be a good option, because the quality of the design-based estimates of the block-means and their sampling variances are independent of model-assumptions on the spatial variation. However, if the spatial variation within the blocks is considerable, then the precision of the estimated block-means can be rather low. If a higher precision is required, then a model-based approach can be considered. It requires quite a few sampling locations, say more than 100–150, in order to obtain a reliable model of the spatial variation (variogram) (Webster and Oliver 1992). Using this model in block kriging might give more precise estimates of the spatial means of the blocks. A design-based alternative for situations where the number of sampling locations per block is rather small, is the synthetic or regression estimator, see Section 4.3.1.1. For more details on the choice between a design-based or model-based sampling strategy, I refer to Brus and de Gruijter (1997).

4.3.1 Design-Based Approach

If a design-based approach is chosen, in each block several locations are selected by probability sampling. If a map is available with units that are related to the contaminant concentrations, i.e. the mean concentration differs between the map units, then we can increase the precision by using the map units as strata in random sampling, see Section 4.2.1.2. Otherwise, random grid sampling within the blocks (Section 4.2.1.3), or stratified random sampling with compact geographical strata within the blocks (Section 4.2.1.2) is a good alternative.

To estimate the mean of a given block, only the data that originate from that block are used. I refer to Section 4.2.2.1 for the estimators of the mean and their sampling variance.

To save laboratory costs, the soil aliquots collected at the sampling locations within a given block can be bulked, see Section 4.2.4. If in decision-making we want to account for uncertainty in the estimated block-means, at least two composite samples must be collected from each block, and separately analyzed.

4.3.1.1 Using Data from Outside the Block

If we can not afford a reasonable number of sampling locations per block so that the quality of the design-based estimates of the block-means will be too poor, then either a model-based sampling strategy can be considered (Section 4.3.2), or a design-based sampling strategy with alternative design-based estimators that make use of the data from outside the block to be estimated. The latter strategy may be considered if the total number of sampling locations is too small to obtain a reliable variogram. The intention is that, by using data from outside the block, the precision of the estimated block mean will increase. On the other hand, in general some bias will be introduced.

Let us consider the situation where the study area can be divided into more or less homogeneous sub-areas (e.g. map units) that cut across the blocks. The sub-areas are used as strata in stratified simple random sampling. Note that we do not have control of the numbers of sampling locations per block in this way. A first estimator of the block-mean, referred to as the synthetic estimator, is

$$\hat{\bar{y}}_{\mathrm{syn},u} = \sum_{h=1}^{L} a_{hu}\,\bar{y}_{\mathrm{s}_h} \ , \tag{4.32}$$

where a_{hu} is the surface area of stratum h in block u relative to the area of the block ($a_{hu} = A_{hu}/A_u$), and $\bar{y}_{\mathrm{s}_h}$ is the sample mean for stratum h. In general the synthetic estimator is not unbiased, but we hope that the reduction of the sampling variance, due to the use of sample data from similar parts of the area outside the block, outweighs the bias.

The second estimator, the regression estimator, tries to eliminate the bias of the synthetic estimator

$$\hat{\bar{y}}_{\mathrm{pos},u} = \sum_{h=1}^{L} a_{hu}\,\bar{y}_{\mathrm{s}_h} + \frac{1}{\widehat{A}_u}\sum_{h=1}^{L}\widehat{A}_{hu}\left(\bar{y}_{\mathrm{s}_{hu}} - \bar{y}_{\mathrm{s}_h}\right) \ , \tag{4.33}$$

where $\bar{y}_{\mathrm{s}_h}$ and $\bar{y}_{\mathrm{s}_{hu}}$ are the sample means in stratum h, and in the intersection of stratum h and block u, respectively, and $\widehat{A}_u$ and $\widehat{A}_{hu}$ are the estimated surface areas of block u and of the intersection of stratum h and block u, respectively. Note that even though we know both surface areas, the estimator uses their estimated areas. This is because this gives higher precision. The areas are estimated by

$$\widehat{A}_{hu} = \frac{n_{hu}}{n_h} A_h \ , \tag{4.34}$$

and

$$\widehat{A}_u = \sum_{h=1}^{L} \frac{n_{hu}}{n_h} A_h \ , \tag{4.35}$$

with n_h and n_{hu} the numbers of sampling locations in stratum h and in the intersection of stratum h and block u, respectively. For estimators of the sampling variance of the synthetic and regression estimators, I refer to de Gruijter et al. (2006, pp. 145, 146).

4.3.2 Model-Based Approach

If the number of sampling locations with a laboratory analysis of the contaminant concentration is large enough to estimate a reliable variogram, a model-based sampling strategy for estimating the block-means is recommendable. In this case there is no need for random sampling, and in general it will be suboptimal. Purposive sampling, such that the sampling locations cover the study area optimally, is a good option in this case. This can be achieved by purposive grid sampling or by spatial coverage sampling, see Section 4.4.1. In purposive grid sampling the grid is not placed randomly on the study area, but located such that the grid nodes optimally cover the study area. For square grids in a square area this is achieved by centering the grid. Regular grids can be suboptimal for irregular shaped areas, and when existing sample data at point locations are available. In these situations a spatial coverage, or in case we have prior measurements at point locations, a spatial infill sample can be a good alternative. They can be designed by forming compact geographical strata by the k-means algorithm, as described in Section 4.2.1. However, in this case the sampling locations are not selected randomly from the strata, but the centroids of the strata are used as sampling locations. For more details on sampling patterns for spatial interpolation, see Section 4.4.1.

In a model-based sampling strategy, the spatial means of the blocks are estimated by block kriging, see Section 4.4.2 for a short introduction to kriging and for references to literature on this geostatistical estimation technique.

4.3.3 Required Number of Sampling Locations

For a design-based approach, constraints can be imposed on the sampling variances of the estimated block-means, on the probability of the error in the estimated block-means, or on the error rates in testing hypothesis on the block-means. If prior estimates of the spatial variances within the blocks, or within the intersections of the blocks and the strata, are available, we can estimate the required number of sampling locations for each block with the procedures described in Section 4.2.5.

To decide on the required number of sampling locations of a model-based sampling strategy, we must choose a variogram before sampling starts. Let us consider first estimating the means of square blocks from a square grid sample. There is no simple equation that relates the grid spacing to the variance of the estimated block means (block kriging variance). What can be done, however, is to calculate the block kriging variance (of the block centred on the midpoints of the grid cells) for a range of grid spacings, plot the block kriging variances against the grid spacing, and use this plot inversely to determine, given a constraint on the variance, the maximum grid spacing. This procedure is implemented in the OSSFIM program, developed by McBratney and Webster (1981). I used this program to compute the maximum grid spacing for estimating the block-average of log(Pb) concentrations in the peat soils near Mijdrecht. The 24 stratified random sampling locations of the reconnaissance survey (Fig. 4.2), supplemented by 108 observations in the neighbourhood with similar soils was used to estimate the variogram of log(Pb) (Fig. 4.6). Note that the log(Pb) concentrations show clear spatial structure. Figure 4.7 shows the

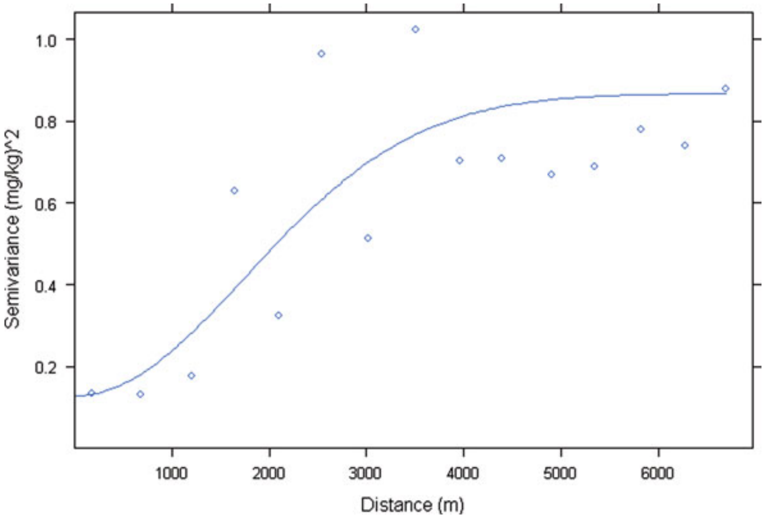


Fig. 4.6 Variogram of log(Pb) concentrations in peat soils at Mijdrecht

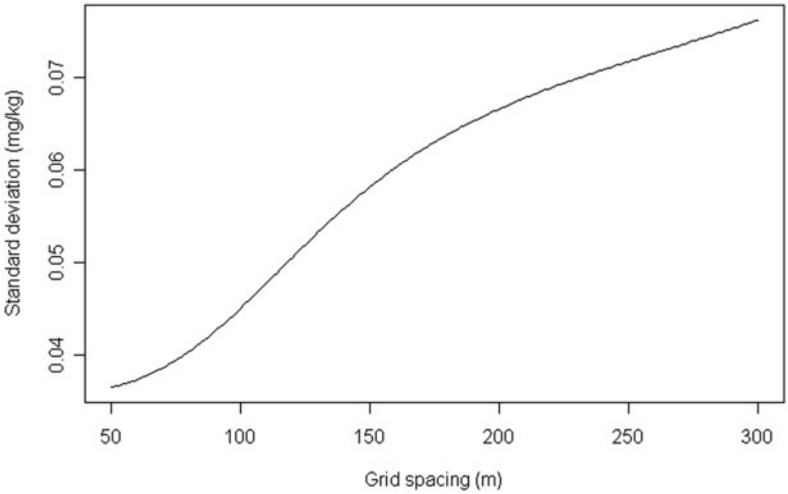


Fig. 4.7 Standard deviation of the predicted average log(Pb) concentration for blocks of 50 × 50 m as a function of the spacing of a square grid

standard deviation of the estimated average log(Pb) concentrations for blocks of 50 × 50 m as a function of the grid-spacing. The plot shows that if, for instance, the standard deviation must be smaller than 0.06 mg kg^{-1}, then the maximum grid spacing is 158 m.

This procedure can also be used to obtain a rough estimate of the required number of sampling locations for spatial coverage sampling, by computing $n = \frac{A}{d^2}$, with A the surface area of the study region, and d the calculated maximum spacing of a square grid. The sampling pattern, i.e. the x and y coordinates of the n sampling locations, is then optimized by k-means clustering or spatial simulated annealing (Section 4.4.1.2). Finally, we may also search for the sampling pattern of n locations with minimum value for the block kriging variance, averaged over all blocks (Section 4.4.1.3).

4.3.3.1 Bayesian Data-Worth Analysis

In Sections 4.2.5 and 4.3.3 the required number of sampling locations is based on a constraint on the uncertainty about the global or local mean. The uncertainty is expressed in terms of variance (sampling variance or block kriging variance), probability of errors in the estimated mean, or probability of decision errors (error rates in statistical testing of hypotheses). The alternative, especially appropriate for sequential sampling, is to base the total number of sampling locations on a cost-benefit analysis. In this approach, the number of new sampling locations, i.e. the number of locations in the next batch, is calculated with the help of a decision model consisting of an objective function for each decision alternative, for instance remediation or no remediation. The objective function includes a risk term, i.e., the extra costs that would be incurred in the event of failure (no remediation of contaminated soil, remediation of clean soil). This risk term is the product of a probability of failure and the costs of failure. By collecting additional sample data, the probability of failure can be reduced. As long as the reduction in the risk outweighs the costs of sampling, sampling is continued. I refer to Freeze et al. (1992); Ramsey et al. (2002); Back (2007); Norberg and Rosén (2006) for more details and applications to survey of contaminated soils.

4.4 Mapping Concentrations at Point Locations

This section is about mapping soil contamination at high spatial resolutions. In the previous section we considered the situation where we want to decide on soil remediation of delineated blocks. In that situation it is natural to take these blocks as estimation units. If the remediation units are not delineated before sampling, but will be delineated after the survey on the basis of the map depicting the contaminant concentration, then in general we would like to have maps depicting mean concentrations of spatial units much smaller than the remediation units. For this aim, a model-based approach is the best (and only) option. This implies that sampling locations need not be selected randomly but typically are selected purposively, and that in estimation (spatial interpolation) a probabilistic model of spatial variation is used. Section 4.4.1 describes various sampling patterns that can be used for spatial interpolation. The next Section 4.4.2 shortly describes how, once the data at the sampling locations are collected, a map of the contaminant concentration can be obtained. The final Section 4.4.3 is about how to determine the required number of sampling locations.

4.4.1 Sampling Patterns

As sampling locations are not selected randomly, we may search for the pattern of sampling locations, i.e. the *x*- and *y*-coordinates, that gives the most precise map. Hereafter I will first describe regular sampling patterns (grids). Next, the constraint of sampling on a grid will be relaxed, generally resulting in irregular patterns. These patterns can be optimized with a criterion defined in terms of distances (between sampling points and the nodes of a fine interpolation grid) leading to spatial coverage and spatial infill samples, or in terms of the estimation error variance, leading to geostatistical samples.

4.4.1.1 Purposive Grid Sampling

Sampling on a regular grid is attractive because of its simplicity. The sampling locations can be positioned in the field easily, especially the nodes of square grids. Contrary to grid sampling in a design-based approach, in a model-based approach the grid need not be placed randomly on the study area. The grid typically is located such that the grid nodes optimally cover the study area. Commonly used grid patterns are square, triangular and hexagonal. Which pattern is optimal depends on the variogram, amongst others. If the contaminant shows moderate to strong spatial autocorrelation, the triangular pattern gives the best result.

Grid sampling is suboptimal when the shape of study area is irregular and when the study area contains enclosures that are inaccessible for sampling, think, for instance, of built-up areas. Shifting the nodes to nearby locations, results in an irregular pattern. Sampling on a grid may also be suboptimal if we want to add new sampling locations to existing ones. The new sampling locations should fill in the empty spaces in between the existing sampling locations.

4.4.1.2 Spatial Coverage and Spatial Infill Sampling

Relaxing the constraint of sampling on a grid, leads to spatial coverage samples, or in case of an additional sample, to spatial infill samples. For such samples, a pattern is calculated that covers the area or fills in the space as uniformly as possible. This is achieved by minimizing a quality measure that is defined in terms of the distances between the nodes of a fine interpolation-grid and the sampling locations. Many distance measures can be selected, see Royle and Nychka (1998). Brus et al. (2007)

proposed to minimize the mean of the squared distances of the grid nodes to their nearest sampling location (Mean Squared Shortest Distance)

$$J_{\mathrm{MSSD}} = \frac{1}{N} \sum_{i=1}^{N} \min_{j} \left(D_{ij}^{2} \right) , \tag{4.36}$$

where n is the total number of nodes of the interpolation grid, and D_{ij} is the distance between the ith grid node and the jth sampling location. This distance measure can be minimized by the simple and fast k-means algorithm. The same algorithm was used in Section 4.2.1.2 to construct compact geographical strata for stratified simple random sampling. In STSI, one or more sampling locations are selected randomly from each cluster of nodes (stratum), whereas here for each cluster the mean x- and mean y coordinate is calculated, and these centroids are used as sampling locations. Whereas in random sampling we may want to have strata of equal surface area (clusters with equal numbers of nodes), so that the sampling design becomes self-weighting, here this constraint should not be used, as it may lead to samples with suboptimal spatial coverage. Existing sampling locations can easily be accommodated in the k-means algorithm, by using them as fixed centroids. Figure 4.8 shows an example.

Alternatively, spatial coverage or spatial infill samples can be designed by the spatial simulated annealing algorithm, as proposed by van Groenigen and Stein (1998). Optimization with spatial simulated annealing requires more skills compared to k-means. There are several parameters in the annealing algorithm that must be chosen by the user and that affect the final sampling pattern. A big advantage of spatial simulated annealing is that it is very flexible with regard to the optimization criterion. For instance, the minimum squared distances in Eq. (4.36) can be weighted to prioritize certain sub-areas, so that the sampling density in these sub-areas is relatively high. This leads to the Weighted Mean Squared Shortest Distance criterion

$$J_{\mathrm{WMSSD}} = \frac{1}{N} \sum_{i=1}^{N} w_i \min_{j} \left(D_{ij}^{2} \right) , \tag{4.37}$$

with w_i the weight attached to node i. van Groenigen et al. (2000) used a similar criterion, Weighted Mean Shortest Distance (note that the distances are not squared), to design a spatial infill sample in an area of ca 30 ha in the old harbour of Rotterdam. Harbour activities are increasingly shifted towards other locations, giving space for house and office building. Existing measurements at 201 locations showed that the concentrations of three heavy metals (Pb, Cu and Zn) and of two carbohydrates (mineral oil and PAH's) frequently exceeded the legal threshold concentrations. To map these contaminants, van Groenigen et al. (2000) designed a first additional sample of 80 locations. A priority map for sampling was made, based on expected contamination and urgency of remediation (Fig. 4.9). The weights attached to the four priority classes were 1.0, 1.5, 2.0 and 3.0. Figure 4.10 shows the pattern of the spatial infill sample. 46 of the 80 locations are located in the most urgent priority

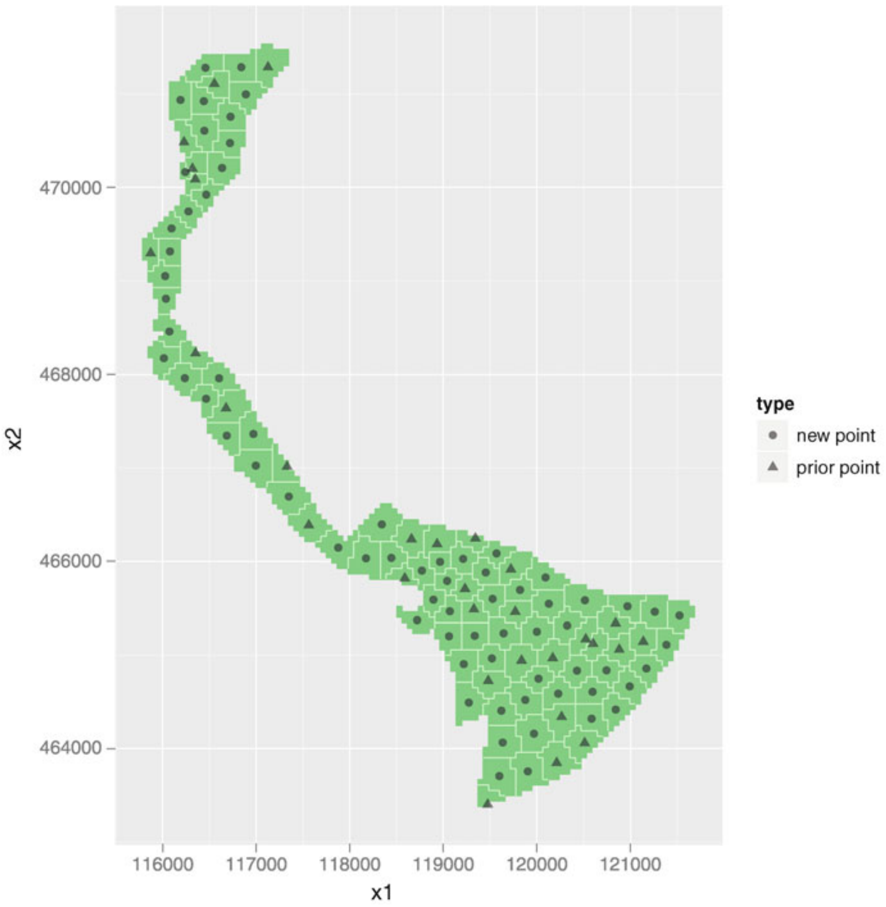


Fig. 4.8 Spatial infill sample of 100 locations obtained by minimizing the Mean Squared Shortest Distance with the k-means clustering algorithm, for mapping heavy metal concentrations at Mijdrecht. The prior sampling locations are the stratified simple random sampling locations of the reconnaissance survey for estimating the spatial cumulative distribution function, see Fig. 4.2. Infill sample computed by R-package spcosa (Walvoort et al. 2009, 2010)

area. Based on these 126 sampling locations, for each of the five contaminants a map of the exceedance probabilities of a specified threshold concentration was computed with indicator kriging. A second infill sample of 30 locations was designed, using the maximum of the exceedance probabilities per contaminant as weights (Fig. 4.11). Several new sampling locations are located close to heavy contaminated locations (Fig. 4.12).

4.4.1.3 Geostatistical Sampling

In the previous section the sampling pattern was optimized in terms of a distance-criterion. Alternatively, we may search for the sampling pattern that minimizes the

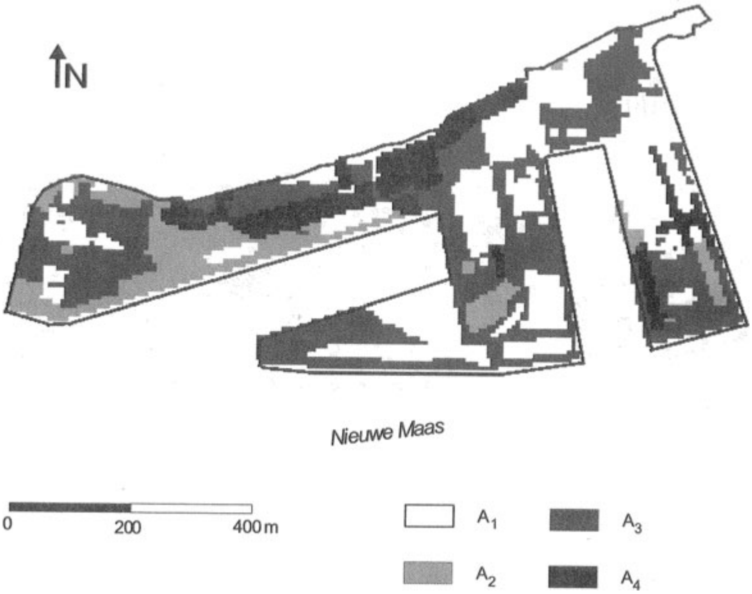


Fig. 4.9 Map with priorities for sampling. A1: lowest priority, A4: highest priority

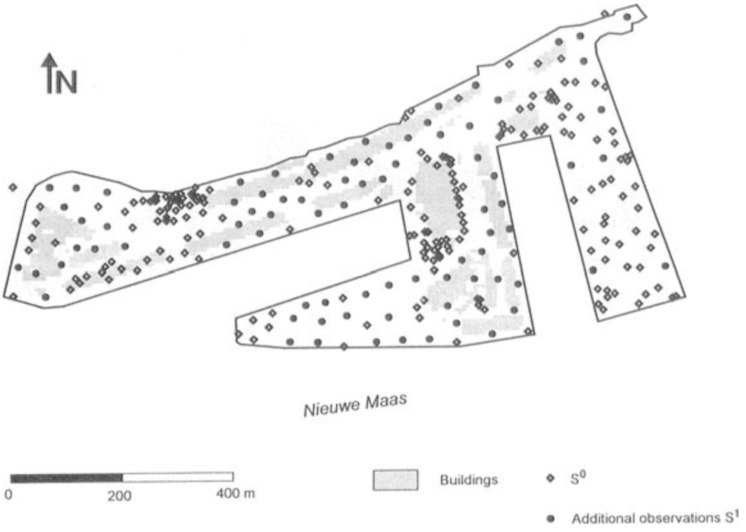


Fig. 4.10 First spatial infill sample (S^1) of 80 locations. S^0: prior sampling locations

error in the interpolations. Evidently, the interpolation errors are unknown, but what can be done is to minimize the variance of the errors. The advantage of interpolation by kriging is that it gives an estimate of this variance, referred to as the kriging variance. The nice thing about kriging is that the variance of the interpolation error is

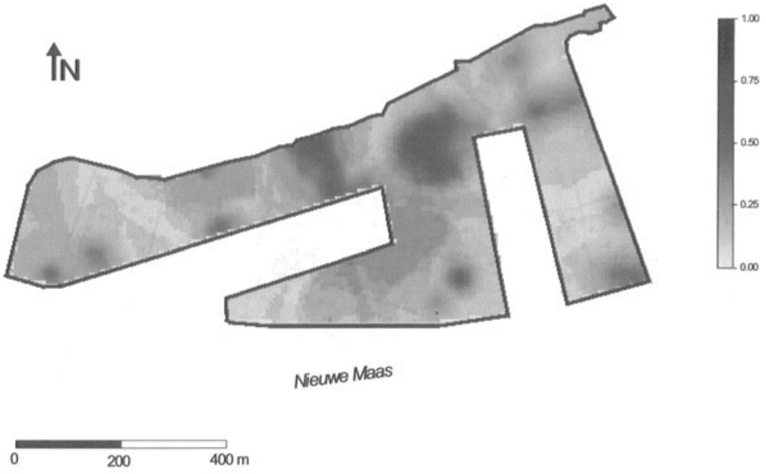


Fig. 4.11 Map with maximum exceedance probability of the threshold concentration, used in selecting second infill sample

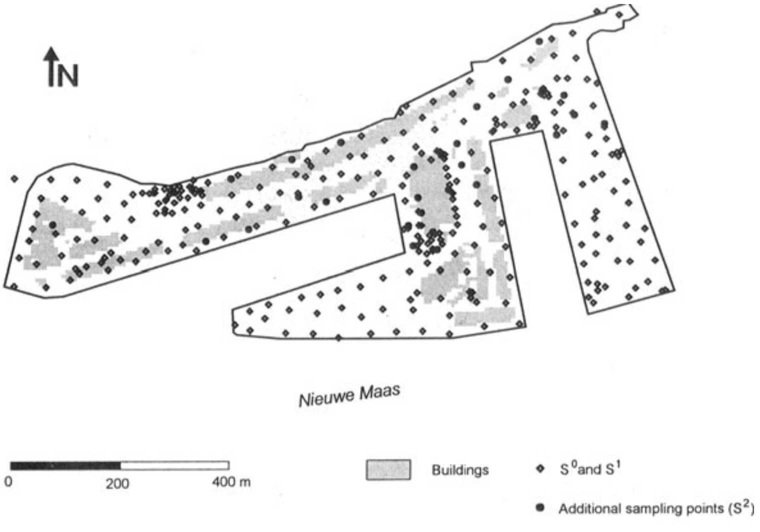


Fig. 4.12 Second spatial infill sample (S^2) of 30 locations

independent of the measured concentrations. So, we do not need the measurements in order to compute the error variance at the nodes of the interpolation grid. This implies that, given the variogram, we may search for the sampling pattern with minimum value for the kriging variance, averaged over the nodes of the interpolation grid

$$J_{\mathrm{MeanV}} = \frac{1}{N}\sum_{i=1}^{N}\left\{V_{\mathrm{K}}\left(\widetilde{Y}_{0i} - Y_{0i}\right)\right\} \ , \tag{4.38}$$

where $V_{\mathrm{K}}\left(\widetilde{Y}_{0i} - Y_{0i}\right)$ is the kriging variance at interpolation node i. Alternatively, the maximum of the kriging variances at the grid nodes may be minimized. To find the sampling pattern with minimum value for Eq. (4.38), van Groenigen and Stein (1998) and van Groenigen et al. (1999) proposed the spatial simulated annealing algorithm. The kriging variance holds for the variogram used in kriging; if we change the variogram, this will affect the kriging variance. Consequently, the optimized sampling pattern also depends on the chosen variogram, although the effect may be rather small.

The problem is that many contaminant concentrations often have distributions that deviate from normal, for instance, the right tail of the distribution can be longer (positive skew or right-skewed). In this case kriging the untransformed concentrations is suboptimal, and more accurate predictions can be obtained by transforming the concentrations such that the transformed concentrations are approximately normal distributed. For this reason concentrations with positive skew are often log-transformed before interpolation by kriging. Using a variogram for the log-transformed concentrations, we may optimize the sampling pattern by minimizing Eq. (4.38). However, it must be stressed that the optimized sampling pattern obtained with the variogram of log-transformed concentrations generally will be suboptimal for the mean variance of the interpolation errors on the original scale.

Brus et al. (2007) found that the difference in the mean kriging variances, Eq. (4.38), of a spatial coverage sample designed with the k-means clustering algorithm and of a geostatistical sample designed by directly minimizing the mean kriging variance with spatial simulated annealing, was small. So, in practice, a useful procedure might be to optimize the sampling pattern with the k-means algorithm first. If this is unsatisfactory for one of the reasons mentioned in Section 4.4.1.2, we might proceed with optimizing the sampling pattern with spatial simulated annealing, either using a distance-criterion such as Eq. (4.36) or, if a prior variogram is available, a variance-criterion such as Eq. (4.38).

A different situation is when one or more maps of covariates are available, think for instance of a map depicting clay content or a digital elevation model. It is well known that contaminant concentrations are often correlated with the clay content or, in floodplains and terraces along rivers, with altitude. In this case interpolation by universal kriging can be advantageous. The optimal sampling pattern for universal kriging generally differs considerably from the optimal pattern for ordinary kriging. The optimal sampling pattern for universal kriging will contain several locations with extreme values for the covariates, enhancing the estimation of the spatial trend parameters (Brus and Heuvelink 2007).

4.4.1.4 Supplementary Sample for Estimating the Variogram

In practice the sample data are used both for estimating the variogram, and for spatial interpolation. In the previous section on geostatistical sampling, it was assumed that the variogram is known. The optimized sampling pattern is optimal for spatial interpolation with this prior variogram. In general this sampling pattern is far from optimal for estimating the variogram. To estimate the nugget (intercept) of the variogram and to choose between alternative variogram model types, for instance, quite a few pairs of points are needed with short mutual distances. Such pairs are missing in grids, spatial coverage samples and geostatistical samples designed for spatial interpolation only. This suggests that some compromise between uniform spatial coverage (for interpolation) and spatial clustering (for estimating the variogram) might be a good option. A simple, practical procedure is

1. split the total sample in two, $n = n_{\text{int}} + n_{\text{var}}$;
2. select purposively n_{int} locations on a grid, or design a spatial coverage or geostatistical sample of n_{int} locations;
3. supplement this first sample by n_{var} locations at short distance of the locations of the first sample;

To avoid spatial clustering of short distance locations in certain parts of the area, we recommend selecting the grid nodes that will receive a short distance location purposively and not at random, for instance by subsampling the regular grid systematically (see Fig. 4.13). Also, we recommend placing the additional locations on the sides of the grid cells, so that the directions for the smallest lag coincide with those of the larger lags.

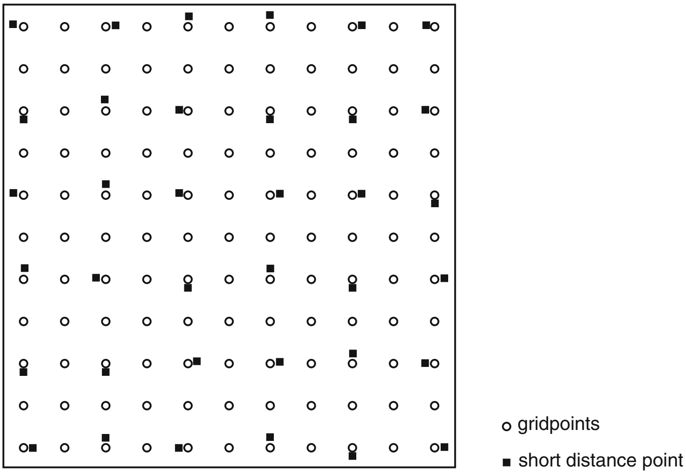


Fig. 4.13 Square grid sample with additional short distance locations for variogram estimation

4.4.2 Spatial Interpolation

Contrary to design-based sampling strategies in which the estimation is based on the method used for selecting the sampling locations (sampling design), in model-based sampling strategies, spatial interpolation is based on the model of spatial variation.

The concentrations at the nodes of the interpolation grid can be estimated by point kriging. There are several versions of kriging, building on slightly different models of the spatial variation, see Goovaerts (1997). The most common version is "ordinary kriging" (OK). In OK it is assumed that the expectation of the concentration is the same everywhere ("stationarity in the mean"). In practice not all sample data are used to estimate the concentration at a given node, but only the sample data in some predefined neighbourhood. This implies that the stationarity assumption is relaxed to the often more realistic assumption of a constant mean within neighbourhoods. In kriging, the concentration at an interpolation node, $\widetilde{Y}_0$, is estimated as a weighted average of the concentrations measured at the sampling locations within the neighbourhood

$$\widetilde{Y}_0 = \sum_{i=1}^{n} \lambda_i \, Y_i \, , \tag{4.39}$$

where Y_i is the measured concentration at the ith sampling location, and λ_i is the weight attached to this location. The weights should be related to the strength of correlation of the concentrations at the sampling location and the interpolation node. The stronger this autocorrelation ("auto" refers to the fact that we consider the same variable at both locations), the larger the weight must be. So if we have a model for this autocorrelation, then we can use this model to find the optimal weights. Usually, not an autocorrelation model is used, but a variogram, which is a model of the dissimilarity of the concentrations at two locations as a function of the distance between the two locations, see Fig. 4.6 for an example. The smaller the semivariance of the concentrations at the interpolation node and the sampling location, the larger the weight must be. Further, if two sampling locations are very close, the weight attached to these two locations should not be twice the weight attached to a single, isolated sampling location at the same distance of the interpolation node. This explains that in computing the kriging weights, besides the semivariances of the n pairs of interpolation node and sampling location, also the semivariances of the $n \cdot (n-1)/2$ pairs that can be formed with the n sampling points are used. For OK, the optimal weights, i.e., the weights that lead to the estimator with minimum error variance (Best Linear Unbiased Estimator), can be found by solving the following $(n+1)$ equations

$$\begin{aligned} &\sum_{j=1}^{n} \lambda_j \, \gamma(h_{ij}) + \nu = \gamma(h_{i0}), \quad i = 1, \ldots, n, \\ &\sum_{i=1}^{n} \lambda_i = 1, \end{aligned} \tag{4.40}$$

where $\gamma(h_{ij})$ is the semivariance of the ith and jth sampling location separated by distance h_{ij}, $\gamma(h_{i0})$ the semivariance of the ith sampling location and the interpolation node separated by distance h_{i0}, and ν an extra parameter to be estimated, referred to as the Lagrange multiplier. This Lagrange multiplier is needed as the error variance is minimized under the constraint that the kriging weights sum to 1, see the final equation in Eq. (4.40). This extra constraint makes the OK-estimator unbiased.

It is beyond the scope of this chapter to explain spatial interpolation by kriging in detail. I refer to Isaaks and Srivastava (1989), for an introduction in geostatistics, to Goovaerts (1997) for an exposé of the many versions of kriging, and to Deutsch and Journel (1997) for kriging software.

4.4.3 Required Number of Sampling Locations

The procedure used to determine the maximum grid spacing for kriging the mean concentrations of blocks, given a constraint on the block kriging variance (Section 4.3.3), can also be used to determine the maximum grid spacing for estimating the concentrations at point locations, given a constraint on the maximum point-kriging variance, see Fig. 4.14 for an example. Note the large standard deviations compared to those of the stimated block-averages, which is due to the averaging out of spatial variation within blocks. As we are uncertain about the variogram, I recommend repeating the calculations with several variograms, and choosing the smallest maximum grid spacing computed with these variograms.

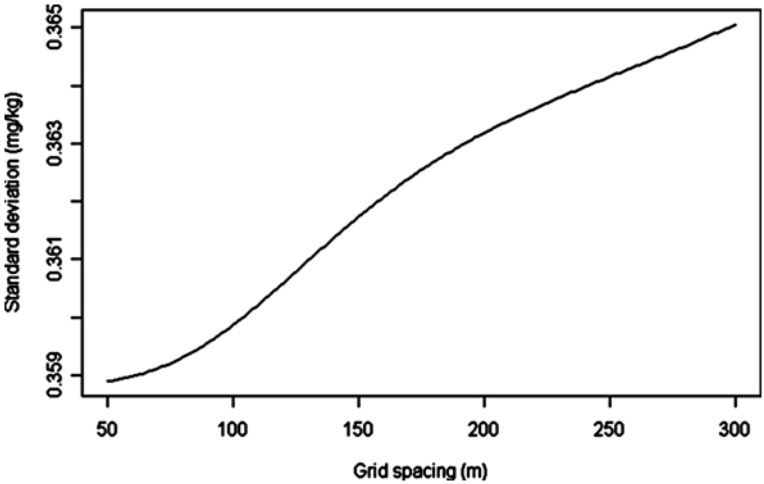


Fig. 4.14 Standard deviation of the estimated log(Pb) concentration at the central point locations of square grid cells as a function of the grid-spacing

4.5 Detecting and Delineating Hot Spots

A hot spot is defined here as a relatively small area with a concentration of a contaminant that exceeds a certain threshold concentration. Examples are sites with soil or groundwater contaminated by some point-source and anomalous sites with high concentrations of natural geochemical elements. The threshold concentration may be a constant value that does not vary in space, or a spatially varying concentration. An example of the latter is a soil standard defined in terms of soil organic-matter content, clay percentage and pH, thus taking into account the bio-availability of the contaminant.

In *detecting* hot spots, the aim of the survey is to find out whether at any point in the study area the critical threshold concentration is exceeded. Finding the exact location is not part of the aim (Section 4.5.1). Once we know that there are locations where the threshold concentration is exceeded, in general the reconnaissance survey is followed by a survey with the aim of *delineating* hot spots, i.e., making a map depicting where the threshold is exceeded, see Section 4.5.2.

A different aim, related to detection of hot spots, is the estimation of the fraction of the area with values exceeding the threshold concentration. For this aim the design-based adaptive cluster sampling strategy can be a good choice (Section 4.2.1.4). An unbiased estimate of the sampling variance of the estimated areal fraction can then be obtained from the sample.

4.5.1 Detecting Hot Spots

The detection of hot spots can be achieved better with purposive sampling than with probability sampling. If one has no prior information on the location of the hot spots, samples are typically taken on a purposively selected, regular grid. Gilbert (1987) worked out a method for calculating the required grid spacing from the consumer's risk, β, i.e., the probability of not hitting a hot spot *if it exists*, and the geometry (size and shape) of the hot spot. This method is implemented in the Visual Sample Plan software (Matzke et al. 2007). The probability of hitting a hot spot if it exists is calculated by summing the zones of coverage for the sampling locations, excluding overlaps. The zone of coverage for any sampling location can be obtained by drawing the contour of a hot spot with its centre at the sampling location. If the centre of the hot spot is in the zone of coverage, it will be detected from the sampling location. In practice, either an elliptical or a circular shape is assumed, and a decision is being made on the minimum size (length of semi-major axis) of any hot spot that we want to detect with a specified probability (Matzke et al. 2007).

So far, it has been assumed that a hot spot exists. In other words, it is assumed that the probability that a hot spot exists is 1. If the existence of a hot spot is uncertain, the probability that a hot spot exists and is detected can be estimated by

$$P(A,B) = P(B|A)\,P(A), \tag{4.41}$$

where $P(B|A)$ is the probability that the hot spot is hit, conditional on its existence, and $P(A)$ is the probability that the hot spot exists. Given the grid spacing and geometry of the hot spot one can calculate $P(B|A)$ and simply multiply this by the a priori probability that the hot spot exists to obtain the probability that the hot spot exists and is detected by the sample.

In some situations, information on land use, or a walkover survey of visual or organoleptic indicators of high concentrations can be used to subdivide the site into sub-areas with different probabilities of containing the hot spot. The grid spacing can then be adapted to these a priori probabilities as follows. For all sub-areas, there must be an equal probability that a hot spot exists when none is detected by the sample. This probability is referred to as the a posteriori probability and denoted by $P(A|\overline{B})$. Bayes' formula can now be used to calculate from the a priori and a posteriori probabilities for each sub-area the probability of not hitting the hot spot if it exists, the consumer's risk β is

$$\beta = \frac{1 - P(A)}{P(A)\left\{\frac{1}{P(A|\overline{B})} - 1\right\}} \ . \tag{4.42}$$

Hence, when *P(A)* differs between subregions, and given a constant $P(A|\overline{B})$ for all sub-areas, for instance 0.05, β differs between sub-areas, and this leads to different grid spacings. Sub-areas with large a priori probabilities will be given smaller grid spacings than sub-areas with small a priori probabilities. An alternative to purposive grid sampling is to optimize the sampling pattern by minimizing the sum of the a priori probabilities outside the zones of coverage of the sampling locations (Tucker et al. 1996).

4.5.1.1 Adding Sampling Locations to the Grid

If none of the grid data exceed the threshold concentration, one may want to add new sampling locations to become more certain about the presence or absence of hot spots. A practical method for selecting new sampling locations is to transform the data so that the univariate distribution is approximately Gaussian, then krige the transformed data, and select the n locations with the smallest values for

$$\zeta_0 = \frac{y_{\mathrm{t}} - \widetilde{Y}_0}{\sqrt{V_{\mathrm{OK}}\{\widetilde{Y}_0 - Y_0\}}} \ , \tag{4.43}$$

where $\widetilde{Y}_0$ is the estimated value at the new sampling location obtained by ordinary kriging, and $V_{\mathrm{OK}}\{\widetilde{Z}_0 - Z_0\}$ is the ordinary kriging variance (Watson and Barnes 1995). Equation (4.43) shows that, depending on the threshold concentration, locations are selected either near sampling locations with a large estimated concentration ($\widetilde{Y}_0$ is large), or in the empty space, where $V_{\mathrm{OK}}\{\widetilde{Y}_0 - Y_0\}$ is large.

4.5.2 Delineating Hot Spots

When no prior measurements on the target variable are available, and the locations of the hot spots are unknown, the best option is purposive grid sampling, or alternatively according to some spatial coverage or geostatistical sample (Section 4.3.2). There are two ways to increase the efficiency of these samples in order to delineate hot spots: sampling in two or more phases (Section 4.5.2.1) and composite sampling (Section 4.5.2.2).

4.5.2.1 Phased Sampling

The efficiency of sampling can be increased by sampling in two, or even more phases (batches). For instance, the first phase could involve sampling the area on a purposively selected square grid. The sample data are then used to design an additional sample of n locations, by estimating the concentration at the nodes of a fine interpolation grid and the kriging variance of these interpolations. Englund and Heravi (1994) proposed to select additional locations by assuming a triangular probability distribution on the interval $[\widetilde{Y} - 3V_{\mathrm{K}}(\widetilde{Y}),\ \widetilde{Y} + 3V_{\mathrm{K}}(\widetilde{Y})]$. This probability distribution is used to calculate for each grid node the probability of decision errors and the expected loss. The node with the highest expected loss is selected as the first additional sampling location. This procedure is repeated (the kriging variances must be updated in each iteration) until the predetermined number of additional sampling locations has been selected.

If the costs of sampling and measurement per sampling location are substantial compared to the costs of decision errors, then the optimal number of sampling locations can be calculated by including this cost component in the loss function. To profit from phased sampling, the number of sampling locations in the first phase must be approximately 75% of the total number of sampling locations (Englund and Heravi 1994).

4.5.2.2 Composite Sampling

Composite sampling is recommended when the laboratory measurement costs are high. Figure 4.15 shows an example where the four aliquots at the corners of square cells are combined to form a composite sample. Analysing the composite sample instead of the individual aliquots reduces the measurement costs, so that more locations can be sampled for the same budget. Due to the larger number of sampling locations, the sample has a better spatial coverage, so that the probability of hitting a hot spot is higher. The problem is that mixing the aliquots implies a loss of information on the concentration in the individual aliquots. The concentration of an individual aliquot may exceed the threshold concentration, while that of the composite sample does not.

Several methods have been developed to identify the individual aliquots (further briefly referred to as “individuals”) with the highest values or with values above

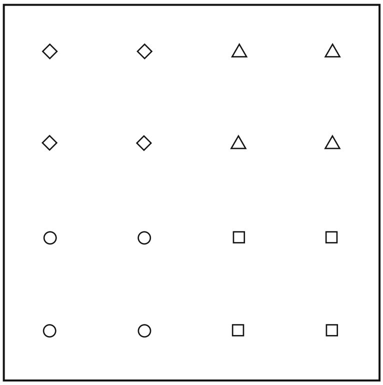


Fig. 4.15 Example of the formation of composites from aliquots taken on a grid. Aliquots from locations with the same symbol form a composite sample

a threshold concentration, by measuring some or all individuals of some composite samples, assuming that some of the soil material of the individuals is kept for later analysis (Gore and Patil 1994, Gore et al. 1996). In these methods, the composite samples are sorted on the basis of their concentration in descending order. For each composite sample, an upper bound of the maximum value is calculated by assuming that the concentration of all except one of the individuals is zero. Under this assumption, the maximum equals the total of the composite sample, i.e., the composite sample concentration multiplied by the number of individuals in that composite sample.

In the simple sweep-out method (Gore and Patil 1994), all individuals of the first ordered composite sample are measured, and the largest value in this composite sample is identified and recorded as the tentative maximum of the individuals of all composite samples (global maximum). If the upper bound of the maximum of the next ordered composite is smaller than this tentative global maximum, then it is clear that this is the true global maximum, and no further composites need to be measured. If the upper bound of the maximum for the next ordered composite is larger than the tentative global maximum, then this composite may contain an individual with a value larger than the tentative global maximum, and the individuals of this composite have to be measured. If the maximum identified in this second composite is larger than the maximum identified in the first composite, then the tentative global maximum is updated. This procedure is continued until the tentative global maximum is larger than the upper bound of the maximum in the next ordered composite.

If the aim is to identify the two largest values, the second largest value is also identified each time the individuals constituting a composite are measured, and the procedure is continued until the tentative second largest value is smaller than the upper bound of the maximum value in the next ordered composite. If the aim is to identify all individuals with values larger than a threshold concentration, then the procedure goes on until the individual with the largest value below the threshold concentration has been identified.

Contrary to the simple sweep-out method, where, once a composite is selected, all individuals except one constituting this composite are measured, the sequential

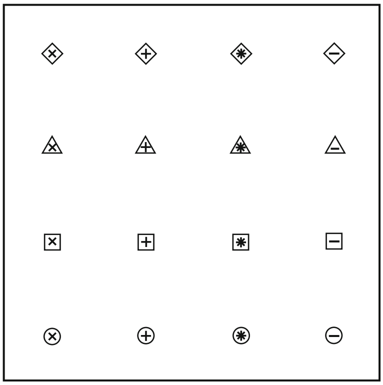


Fig. 4.16 Example of two-way composite sampling on a square grid. At the nodes of the grid, two soil aliquots are collected, one for a row composite sample, one for a column composite sample

sweep-out methods (Gore et al. 1996) require measuring individuals of a composite as long as there is a possibility that there is a larger value among the unmeasured individuals of that composite. After each analysis of an individual, the total of the remaining, unmeasured individual samples is updated by subtracting the concentration of the measured individual sample from the previous total. This updated total is used as the upper bound of the maximum of the "residual" composite, i.e., the composite that could be formed by mixing the remaining individuals. The procedure stops when the tentative global maximum exceeds the updated upper bound of the maximum of all (residual) composites.

In the sweep-out methods described above, the individual to be measured in a selected composite is selected randomly, which is reasonable because there is no information available on the individual values. When composites are formed by a two-way compositing design some information is available, and it is possible to select the individual with the highest probability of having the maximum (Gore et al. 2001). Figure 4.16 shows an example of a two-way composite sampling design. From every individual aliquot, two portions are taken, one of which is combined with the portions from the same *column* to form a column composite, the other is combined with the portions from the same *row* to form a row composite. The individual that contributes to the column composite with the largest value *and* to the row composite with the largest value has the largest probability of having the maximum value, and will therefore be selected for measurement. For a sweep-out method for this two-way compositing design, I refer to Gore et al (2001).

In the above methods, all values above a threshold concentration are identified with certainty. The alternative is to accept some uncertainty, and to continue measuring individuals until, for all (residual) composites, the probability that they include an individual with a value above the threshold concentration is smaller than some chosen limit (Carson 2001). Correll (2001) proposed to measure the individuals of all composites having a concentration higher than $z_t/\sqrt{k}$, where z_t is the threshold concentration for the individual values, and k is the composite size. This modified threshold concentration gave few false negatives (i.e., cases where the composite is below the modified threshold concentration, while at least one of the individuals exceeds the original threshold concentration), and few false positives (i.e., cases where the composite exceeds the modified threshold concentration, while none of the individuals exceeds the original threshold concentration).

Acknowledgements I would like to thank R.O. Gilbert for his valuable comments on this chapter.

References

Back P-E (2007) A model for estimating the value of sampling programs and the optimal number of samples for contaminated soil. Environ Geol 52:573–585

Boswell MT, Gore SD, Lovison G, Patil GP (1996) Annotated bibliography of composite sampling, part A. Environ Ecol Stat 3:1–49

Brus DJ, de Gruijter JJ (1997) Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma 80: 1–59

Brus DJ, de Gruijter JJ (2003) A method to combine non-probability sample data with probability sample data in estimating spatial means of environmental variables. Environ Monitor Assess 83:303–317

Brus DJ, de Gruijter JJ, van Groenigen JW (2007) Digital soil mapping. An introductory perspective, chapter 14: Designing spatial coverage samples using the k-means clustering algorithm. Elsevier, Amsterdam, pp 183–192

Brus DJ, Heuvelink GBM (2007) Optimization of sample patterns for universal kriging of environmental variables. Geoderma 138:86–95

Brus DJ, Lame FPJ, Nieuwenhuis R (2009) National baseline survey of soil quality in the Netherlands. environmental pollution. Environ Pollut 157:2043–2052

Brus DJ, Nieuwenhuizen W, Koomen A (2006) Can we gain precision by sampling with probabilities proportional to size in surveying recent landscape changes in the Netherlands? Environ Monitor Assess 122:153–169

Brus DJ, Spätjens LEEM, de Gruijter JJ (1999) A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. Geoderma 89:129–148

Carson JH (2001) Analysis of composite sampling data using the principle of maximum entropy. Environ Ecol Stat 8:201–211

Chang Y-C, Yeh H-D (2007) Optimum allocation for soil contamination investigations in Hsinchu, Taiwan, by double sampling. Soil Sci Soc Am J 71:1585–1592

Cochran WG (1977) Sampling techniques. Wiley, New York

Correll RL (2001) The use of composite sampling in contaminated sites – a case study. Environ Ecol Stat 8:185–200

de Gruijter JJ, Brus DJ, Bierkens MFP, Knotters M (2006) Sampling for natural resource monitoring. Springer, Berlin, Heidelberg, New York

Deutsch CV, Journel AG (1997) *GSLIB*. Geostatistical software library and user's guide, 2nd ed. Oxford University, Oxford

Elder RS, Thompson WO, Myers RH (1980) Properties of composite sampling procedures. Technometrics 22:179–186

EPA (2006) Guidance on systematic planning using the Data Quality Objectives Process. Technical Document EPA QA/G-4, EPA/240/B-06/001, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, DC

Englund EJ, Heravi N (1994) Phased sampling for soil remediation. Environ Ecol Stat 1:247–263

Freeze RA, James B, Massmann J, Sperling T, Smith L (1992) Hydrogeological decision analysis: 4. The concept of data worth and its use in the development of site investigation strategies. Ground Water 30:574–588

Gilbert RO (1987) Statistical methods for environment pollution monitoring. Van Nostrand Reinhold, New York

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University, New York

Gore SD, Patil GP (1994) Identifying extremely large values using composite sample data. Environ Ecol Stat 1:227–245

Gore SD, Patil GP, Taillie C (1996) Identification of the largest individual sample value using composite sample data and certain modifications of the sweep-out method. Environ Ecol Stat 3:219–234

Gore SD, Patil GP, Taillie C (2001). Identifying the largest individual sample value from a two-way composite sampling design. Environ Ecol Stat 8:151–162

Isaaks EH, Srivastava RH (1989) An introduction to applied geostatistics. Oxford University, New York

Lohr SL (1999) Sampling: design and analysis. Duxbury, Pacific Grove, USA

Matzke B, Wilson J, Nuffer L, Dowson S, Gilbert R, Hassig N, Hathaway J, Murray D, Sego L, Pulsipher B, Roberts B, McKenna S (2007) Visual sample plan version 5.0 users guide. Technical Report PNNL-16939, Pacific Northwest National Laboratory, Richland, WA, US

McBratney AB, Webster R (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables: II Program and examples. Computers Geosci 7:335–365

Nanthakumar A, Selvavel K (2004) Estimation of proportion of success from a stratified population: a comparative study. Commun Stat 33:2245–2257

Norberg T, Rosén L (2006) Calculating the optimal number of contaminant samples by means of data worth analysis. Environmetrics 17:705–709

Ramsey MH, Taylor PD, Lee J (2002) Optimized contaminated land investigation at minimum oveall cost to achieve fitness-for-purpose. J Environ Monitor 4:809–814

Rohde CA (1976) Composite sampling. Biometrics 32:273–282

Royle JA, Nychka D (1998) An algorithm for the construction of spatial coverage designs with implementation in SPLUS. Computers Geosci 24:479–488

Särndal CE, Swensson B, Wretman J (1992) Model assisted survey sampling. Springer, New York

Thompson SK, Seber GAF (1996) Adaptive sampling. Wiley, New York

Tucker P, Ferguson CC, Abbachi A, Nathanai P (1996) A decision support system for designing sampling strategies for potentially contaminated sites. In: Kraak MJ, Molenaar M (eds) Spatial data handling. Proceedings of the seventh international symposium on spatial data handling, Delft, The Netherlands, August 1996, vol 1. International geographical union commission on geographic information systems, Association for geographical information, pp 1–12

van Groenigen JW, Pieters G, Stein A (2000) Optimizing spatial sampling for multivariate contamination in urban areas. Environmetrics 11:227–244

van Groenigen JW, Stein A (1998). Constrained optimization of spatial sampling using continuous simulated annealing. J Environ Qual 27:1078–1086

van Groenigen JW, Siderius W, Stein A (1999) Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma 87:239–259

Walvoort D, Brus D, de Gruijter J (2009) spcosa: Spatial coverage sampling. Avalilable at cran.r-project.org

Walvoort D, Brus D, de Gruijter J (2010) An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Comput Geosci. doi:10.1016/j.cayeo2010.04.005

Watson AG, Barnes RJ (1995) Infill sampling criteria to locate extremes. Math Geol 27:589–608

Webster R, Oliver MA (1992) Sample adequately to estimate variograms of soil properties. J Soil Sci 43:177–192

Wendell JP, Schmee J (1996) Exact inference for proportions from a stratified finite population. J Am Stat Assoc 91:825–830

Part III
Human Health Aspects

Chapter 5
Human Health Risk Assessment

Frank A. Swartjes and Christa Cornelis

Abstract Exposure of humans to contaminated sites may result in many types of health damage ranging from relatively innocent symptoms such as skin eruption or nausea, on up to cancer or even death. Human health protection is considered as a major protection target, both by decision-makers as well as by the general public. The first step in Human Health Risk Assessment is definition of the problem (issue framing). In this stage, the scope of Human Health Risk Assessment must be clearly defined and the various stakeholders need to be actively involved. It is important to define the timeframe for which the Risk Assessment is applicable, since the effects depend on the duration of exposure and factors that impact human health risk will change over time. Subsequently, Exposure Assessment and Hazard Assessment must be performed. Ideally, the Exposure Assessment covers a smart combination of calculations, using exposure models, and measurements in contact media and body liquids and tissue (Biomonitoring). Hazard Assessment, which is different for contaminants with or without threshold effects, results in a Critical Exposure (aka: Toxicological Reference Value). In a final step, Risk Characterisation provides a risk appraisal calculated on the basis of exposure and hazard. Specific attention is given in this chapter to phenomena such as public perception, probabilistic Human Health Risk Assessment, Physiologically-Based PharmacoKinetic modelling, background exposure, sensitivity and uncertainty analyses, human health-based Soil Quality Standards, site-specific Human Health Risk Assessment on the basis of a tiered approach and ethical issues in regard to testing of human beings.

F.A. Swartjes (✉)
National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: frank.swartjes@rivm.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_5,

Contents

5.1 Introduction

5.1.1 Threat to Human Health

The contact of humans with contaminants may result in many types of health damage, ranging from relatively innocent symptoms such as skin eruption or nausea, on up to cancer or even death. Contaminants have widely differing potentials for causing health damage. However, contaminated sites can only lead to an adverse health impact if humans are exposed to the contaminants. This exposure can result from direct contact with the soil, or after transfer of the contaminants to so-called *contact media* (e.g., vegetables, indoor air), and subsequent exposure to these contact media (e.g., due to consumption or inhalation). Exposure to soil contaminants can occur by oral, inhalation and dermal routes, thus relating to the pathway within the human body through which the contaminants enter the body (the mouth, gullet, stomach; the nose, trachea and lungs; and the skin; respectively).

It must be taken into account that, apart from contaminated sites as sources for contaminants, contaminant cycling is part of human life. Humans are surrounded by and in close contact with contaminant-holding materials and matrices, on a daily basis. In fact, the human body is loaded with contaminants (Schroijen et al. 2008; WWF 2003). The chemical load of the human body can be considered as a chemical footprint of past and present human contact with contaminant holding materials, including soil, groundwater and contact media that relate to soil and groundwater. Budd et al. (2004), for example, demonstrated on the basis of measurements in tooth enamel of 77 individuals buried in England that humans were exposed to lead from geological sources via their diet, at least since the Neolithic Era (about 5,000 years ago).

At the overall population level, exposure due to contaminated sites is considered to have a limited impact on pubic health. At the local level, however, exposure to a contaminated site can dominate the exposure to specific contaminants and thus pose a significant risk to human health.

In urban areas, low concentrations of contaminants are often found in soils. Manta et al. (2002), for example, found elevated concentrations of metals in of Zn, Cu and Hg (medians of 138, 63 and 0.68 mg/kg$_{\text{soil, dw}}$, respectively, in the city of Palermo (Sicily), Italy. However, the authors also found a relatively high concentration of lead (median 2002 mg/kg$_{\text{soil, dw}}$), a phenomenon often found in urban areas. Teaf (2008) found relatively low PAH concentration in urban areas in Florida, USA (1–5 mg/kg$_{\text{soil, dw}}$ expressed in benzo(a)pyrene-equivalents). Although those concentrations frequently exceeded the default Florida cleanup target level for both residential and commercial/industrial land use, it was concluded that they were completely consistent with levels reported in a great many urban settings.

5.1.2 Public Perception

Good health is, without any doubt, both literally and figuratively, a priceless asset. Almost every person in this world would probably put good health in the number-one position, when asked what the most important things in life are. Since good health is such a precious thing, not many arguments are needed to convince decision-makers or the general public of the need for protecting human beings from the health effects due to contaminated sites. In this regard and contrary to the protection target 'soil ecology', for example, one does not have to be a scientist to understand the significance of the protection target 'human health'. As a consequence, human health is widely recognised as the major protection target in the risk-based assessment of soil quality and the management of contaminated sites in Europe (Carlon and Swartjes 2007). Although the extent of health damage due to soil contamination is often debated and over-estimated, it is generally accepted that as part of modern civilisation people must be safe on the sites where they live, work, or recreate.

The uncontrollable and often unobservable nature of soil contamination generally results in a higher perceived risk than, for example, risk from smoking or bungee jumping, as these latter are people's conscious choices (International Programme on Chemical Safety 1999). For that reason, humans require the presumed responsible party (e.g., the local authorities or the industry) to invest in contaminated site management and are even willing to contribute as a taxpayer.

5.2 Principles of Human Health Risk Assessment

5.2.1 Problem Definition

As mentioned in Section 1.1.4.2, from a more general Risk Assessment perspective, the first step in a Human Health Risk Assessment is the *definition of the problem* or *issue framing*. For this purpose, the scope of the human health risk-based soil quality assessment must be clear. Human health Risk Assessment for contaminated sites can have a series of objectives, e.g., the derivation of Soil Quality Standards, the assessment of actual or potential human health risks, the development of Remediation Objectives, or the ranking of contaminated sites in regard to the risks for human health.

At the problem-definition stage the various stakeholders need to be involved. Since regulators often have a profound impact on the initiation and performance of a risk-based soil or groundwater quality assessment, it seems appropriate that they formulate the exact purpose of the assessment. Regulators also are responsible for definition of the boundary conditions such as definition of the groups that have to be protected (e.g., adults, children, workers, vulnerable groups such as pregnant women) and the required precaution/conservatism of Risk Assessment. Therefore, an intensive communication between regulators and scientists, and, in fact, with all stakeholders involved, must take place in this early stage of a contaminated site project.

The Risk assessor must be aware of any aspect that influences human health risks such as present land use and, when appropriate, desired future land use, history of the site in so far as this may affect the contamination of the soil or groundwater, soil and groundwater properties, and the facts and figures about human behaviour on the site. It is of crucial importance to have insight into all information about the site, including possible former reports that were written about Risk Assessment-related topics.

5.2.2 Risk Characterisation

As explained in Section 1.5.3, from a more general Risk Assessment perspective, the *Human Health Risk Characterisation* is preceded by two steps, i.e., the Exposure Assessment (representing 'probability' or 'chance') in Risk Assessment terms (risk is the multiplication of probability or chance, and effect) and the Hazard Assessment (representing 'effect' in Risk Assessment terms). Basically, Human Health Risk Assessment is all about linking exposure to effects. Issues that need attention in the Risk-Characterisation phase are the combination of oral, inhalation and dermal exposures, relevant timeframes for exposure in regard to the occurrence of effects, and the compatibility of Estimated Exposure and Critical Exposure (as *Toxicological Reference Value*). The Risk Characterisation is a crucial step in Human Health Risk Assessment, since it results in a judgement about the contaminated status of the site in regard to human health, a widely accepted protection target. Risk Characterisation is described in detail in Section 5.5.

5.2.3 Communication

Terminology in regard to human health risk appraisal can lead to an enormous amount of confusion among the general public. Obviously, phrases such as 'potential exposure does not exceed the Toxicological Reference Value' overreach clear-communication standards and often even arouse suspicion in the layman. The phrase 'there is no risk for human health', when respecting the basic principles of Risk Assessment, may be used only in case of the absolute absence of any soil contaminants. On the other hand, the presence of a risk for human health, even when it is an acceptable risk without any further consequences and/or actions required, if not properly conveyed, might be translated into 'a bad situation that they (the regulators) are trying to mask', when the meaning of this risk is not communicated clearly.

Scientifically, the phrase *tolerable human health risk* is often used. It represents an objective level of risk. Politically, the preferred phrase is *acceptable human health risk*. It represents the politically acceptable level of risk, usually chosen on the basis of scientific information about the range of risk levels. For simple communication purposes the word *safe* may be used, although it must be noted that this is a subjective term.

5.3 Exposure Assessment

5.3.1 Definition

In popular terms, exposure is the amount of a contaminant that enters the human body, expressed as contaminant mass, per unit of body weight and time ($\text{mg.kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$). The more formal IPCS (International Programme on Chemical Safety) definition is: the concentration or amount of a particular agent that reaches the body in a specific frequency for a defined duration (International Programme on Chemical Safety 1999, 2004). This definition implies that exposure could represent a rate, expressed as mass per unit of body weight and time ($\text{mg.kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$), but also as concentration, expressed as mass per unit of air volume (mg/L). Both exposure metrics can be used for comparison with the Toxicological Reference Values, either as a rate or a concentration, in the Risk Characterisation step. In this chapter often the terms *Estimated Exposure* and *Critical Exposure* are used. Estimated Exposure is the best possible estimate of human exposure, either through modelling and/or measuring. Critical Exposure enables a distinction in situations that do or do not need further attention with regard to human health risks, either based on science and/or political decisions.

Ideally, the relevant measure of human exposure is the amount of contaminants that is actually absorbed into the body, i.e., the *internal exposure* (aka: 'dose' or 'uptake'). More precisely, human exposure relates to the contaminants that are adsorbed into specific target organs, that is, the organs in which adverse effects are revealed. Therefore, instead of dose-response relationships, it is more accurate to refer to *exposure-dose-response relationships* (or external exposure-internal exposure-effect relationships).

Exposure Assessment provides a (quantitative) evaluation of exposure, including intensity, frequency and duration of exposure, route of exposure (oral, inhalation, dermal), rates (intake or uptake rates), the amount that may cross the body boundary (external exposure) and the amount absorbed (internal dose) (International Programme on Chemical Safety 1999). However, since the fate of contaminants in the human body is difficult to assess, certainly on the basis of calculations, exposure often refers to the amount of contaminants that cross the outside borders of the human body, that is, the *external* exposure (aka: 'intake'). Since external exposure represents the maximum dose, that is, the dose when 100% of all contaminants have been absorbed into the body, this can be regarded as a worst-case approach.

5.3.2 Biomonitoring

The most direct way to assess human exposure from soil contaminants is to measure the actual body burden through *Biomonitoring*. In practice this means sampling and measuring body fluids or body tissue. For several reasons, however, these measurements often offer limited possibilities and are only used in higher tier Risk Assessments. First of all, sampling of the tissue of living humans is difficult, or impossible, for ethical and technical reasons. For clear reasons it is next to impossible to sample organ tissue of living humans. The best options offer the sampling of the following materials:

- blood, for example, Kim et al. (2005), who measured persistent contaminants such as polybrominated diphenyl ethers (PBDEs) and polychlorinated biphenyls (PCBs) in human blood, in Seoul, Korea;

- urine, for example, Kouniali et al. (2003), for measuring benzene metabolites in urine of adults and their children in France; Standaert et al. (2008), who measured cadmium in urine of adults in Belgium;
- nails, for example, Rodushkin and Axelsson (2003), who measured 55 metal and metalloids in nails of adults in Sweden;
- hair, for example, Pereira et al. (2004), who measured arsenic, cadmium, copper, manganese and zinc in scalp hair of the human population living near an abandoned pyrite mine in Alentejo, Portugal.

In extreme cases, skin tissue may be sampled. Indirect sampling of human material entails the measurement of volatile compounds in exhaled breath (e.g., Miekisch et al. (2004), who measured volatile organic contaminants in exhaled breath in Germany). Walker et al. (2003) based their Risk Assessment on measurements of PCBs and several organochlorine pesticides (such as DDT, DDEs, β-HCH and hexachlorobenzene) in maternal and umbilical cord blood plasma of several ethnic groups in the Canadian Arctic. Signorini et al. (2001) performed, additionally to soil and blood sampling, a health monitoring program of the population, including measurements of the liver function, immune function, neurological impairment, dermatological effects, reproductive pathology, and mortality due to exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin originating from the industrial Seveso accident in Italy, in 1976.

The big advantage of sampling body fluids or body tissue is that this generally results in more reliable data concerning the *actual* body burden at the time of sampling than would any calculation. Although models exist (see Section 5.6.3 on Physiologically-Based PharmacoKinetic modelling), the calculation of internal exposure and the subsequent excretion, metabolism and bioaccumulation is complex and relatively unreliable.

Measuring contaminants in body fluids or body tissue often pretends to be *very* reliable. This may, in turn, lead to overestimation of the value that is given to these measurements. Due to technical difficulties, or spatial and/or temporal variation, measurements do not *per definition* result in more reliable values. For risk-based soil quality assessment purposes, Biomonitoring also has the disadvantage that the measured contaminants in body fluids or body tissue may not only originate from the soil, but also from other sources, for example, the diet. In addition, one needs to account for the time-scale that is represented by the measurement. As an example, cadmium in urine represents a measure of cumulative lifelong exposure, whereas cadmium in blood represents recent exposure. Cadmium in urine thus would reflect a person's exposure history. If exposure had changed over time due to changes in residence conditions or overall environmental quality, an erroneous interpretation of human health risks from the contaminated site could result. In relation to the relevant time-scale, different choices are made for the statistical interpretation of the measured concentrations in body fluids and tissues (Iversen et al. 2003).

When Biomonitoring is considered and adequate methods are available, the objective of the study should be clearly defined. If the results will be used in Human Health Risk Assessment, Toxicological Reference Values should exist for the parameters measured. If the purpose of the assessment relates to prioritizing human health risks in different areas, for example, attention should be paid to the statistical boundary conditions for comparing population groups when interpreting elevated body concentrations of contaminants.

As a negative side effect, sampling body fluids or body tissue might bring about a socially and physiologically adverse impact on humans. Such an impact is hard to quantify.

In practice, sampling of body fluids or body tissue is, although relatively costly, possible and useful for a few contaminants. The most frequently used measurement program relates to the measurement of *blood lead levels* (BLLs) in children, which is representative for exposure during childhood, such as, for example, that performed in Maynard et al. (2003) for sampling of lead in the blood of children in South Australia. In Fig. 5.1, a picture of the performance of a blood lead sampling monitoring program is shown in response to community concerns of possible elevations in blood lead levels among the children of Anniston, Alabama, USA, during spring 2001 (Thompson 2002).

The current Critical Exposure value of lead in blood in children of 10 μg/dL (e.g., Johnson and Bretsch (2002), who used blood concentrations higher than 10 $\mu g\ dL^{-1}$ as the criterion for elevated blood levels at a contaminated site in Syracuse, USA). This value of 10 μg/dL, although presently a subject of debate (e.g., Gilbert and Weiss (2006), who claimed that blood lead levels below 10 $\mu g \cdot dL^{-1}$ may impair neurobehavioral development and argued that a level of 2 $\mu g \cdot dL^{-1}$ is a useful and feasible replacement), offers the possibility for appraisal of the measured blood levels. One critical note about lead-blood sampling, however, is that lead-blood levels are influenced by lead exchange between the blood and the bones, especially by children, and there is debate about the residence time of lead in blood (Mushak 2003).

Fig. 5.1 The performance of a blood lead sampling monitoring program in response to community concerns of possible elevations in blood lead levels among the children of Anniston, Alabama, USA, during spring 2001 (source: ATSDR; reproduced with permission)

In summary, Biomonitoring can be very supportive for site-specific Risk Assessment, mostly in higher tier Human Health Risk Assessment and, generally, for large-scale contamination, since the costs involved are relatively high. However, as mentioned above, these measurements also have some drawbacks. In Section 1.7.3, a more general viewpoint on the benefits and disadvantages of measurements is given. It was concluded that the ideal Human Health Risk Assessment is based on a smart combination of calculations (less accurate, but more representative for the long term) and measurements (more reliable, but mainly representative for the short term).

5.3.3 Exposure Calculations

5.3.3.1 Exposure Models

A very useful and practical methodological possibility for assessing human exposure is that of *calculating* human exposure, using a so-called *exposure model*. These exposure models enable the calculation of the rate of soil contaminants that enter the human body, blood stream, or target organs. Exposure models consider direct contact with the soil and intake of so-called contact media that include contaminants from the soil. Contact media are environmental compartments, which are in direct contact with the soil and to which soil contaminants can migrate, such as vegetables, indoor air and house dust.

Most of the existing exposure models use a representative *total soil concentration* as a starting-point. Depending on the purpose of the Human Health Risk Assessment, this representative soil concentration could be a relatively high estimate of actual concentration (e.g., in a first tier Human Health Risk Assessment) or an average soil concentration.

In Fig. 5.2, the layout of the Dutch exposure model CSOIL (Brand et al. 2007; Van den Berg 1991/1994/1995) is given, as an example. Three elements are recognized in exposure models (see Fig. 5.2):

- The determination of the contaminant distribution over the soil compartments, that is, the solid phase, pore water and soil air.
- The determination of contaminant transfer from (the different mobile compartments of) the soil into contact media (*intercompartimental relationships*).
- The calculation of (direct and indirect) exposure of humans.

Similar often cited exposure models are CalTOX from the USA (McKone 1993), CLEA from the UK (DEFRA and EA 2002) and VlierHumaan from Flanders, Belgium (Provoost et al. 2004a).

Current Risk Assessment models generally limit exposure calculations to the calculation of relevant concentrations, for example, in indoor air, or to *external* exposure. For the majority of soil contaminants, Toxicological Reference Values exist for these exposure metrics, and Risk Characterisation is thus possible.

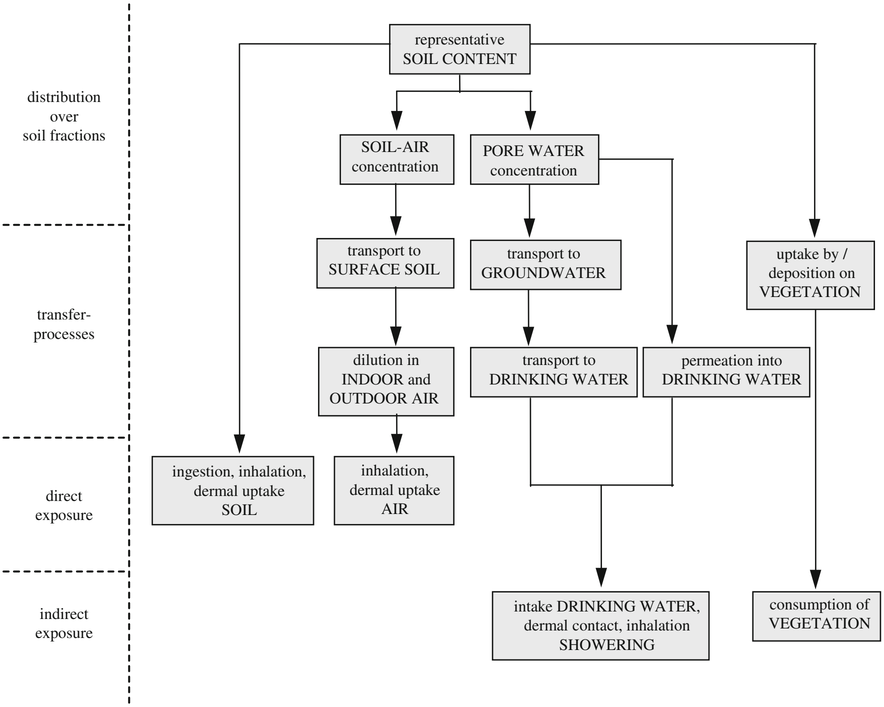


Fig. 5.2 Layout of the CSOIL model, as an example of the layout of exposure models

5.3.3.2 Contaminant Distribution

The distribution of contaminants over the soil phases (solid phase, pore water and soil gas) can be calculated on the basis of partition equations. Partition or distribution coefficients describe the (equilibrium) relationship between the concentrations in the solid soil phase and the pore water, and between the pore water and the soil air, as follows:

$$\rho_b{}^* C_{Tot} = \theta_{pw}{}^* C_{pw} + \rho_b{}^* C_s + \theta_a{}^* C_a \tag{5.1}$$

where

C_{Tot} is the total soil concentration (M/$\mathrm{kg_{dw}}$), C_{pw} (M/$\mathrm{L_{pw}}$), C_{s} (M/$\mathrm{kg_{dw}}$) and C_{a} (M/$\mathrm{L_{air}}$) are the concentrations in pore water, solid soil phase and soil air, respectively, ρ_{b} is the bulk density of the soil (M/L) and θ_w and θ_a represent volumetric water and air content of the soil (–), respectively.

The concentrations in the mobile phases of the soil can be calculated as follows:

$$C_{pw} = C_{Tot}/K_d \tag{5.2}$$

$$C_a = H' * C_{pw} \tag{5.3}$$

where K_d is the distribution coefficient between soil solid phase and pore water (L/M) and H′ (the dimensionless Henry coefficient (–)) is the distribution coefficient between soil air and pore water.

Alternatively, the distribution of contaminants over soil phases can be determined on the basis of the fugacity theory of Mackay (2001). Fugacity is the tendency of a contaminant to escape from the medium in which it is present and has the dimensions of pressure ($\mathrm{M{\cdot}L^{-1}T^{-2}}$). At low concentrations fugacity is proportional to concentration, as follows:

$$C = Z^* f \tag{5.4}$$

where Z is the fugacity capacity (T^2/L^2) and f is fugacity ($M{\cdot}L^{-1}{\cdot}T^{-2}$).

For air, Z equals $1/R^*T$, for water Z equals $1/H$ and for solid phases Z equals $K_d{}^*\rho_b/H$. In a multi-compartmental system at equilibrium, the fugacities are equal and concentrations in the soil compartments can be calculated from the Z values. The fugacity approach was initially applied to organic contaminants distributing over all phases. If the distribution of non-volatile contaminants were to be modelled, modifications to the equations would be needed.

5.3.3.3 Contaminant Transfer

The second step in exposure modelling is the prediction of the transfer of the contaminants from the mobile phase of the soil to contact media, such as:

- air (volatile phase), outdoor and indoor;
- air (particles);
- water (groundwater, surface water, drinking water);
- household dust;
- vegetables;
- meat;
- eggs;
- milk.

Accumulation in cattle can be considered as transfer or as exposure, depending on the protection targets considered (human health or the cattle).

In the final step, contact by humans with the contaminants in these compartments is quantified, taking into account site characteristics and resident properties.

5.3.3.4 Major Exposure Pathways

Formulae for the following exposure pathways have been included in almost all of the existing exposure models (Carlon and Swartjes 2007; McKone 1993):

- exposure through soil ingestion (oral exposure);
- exposure through vegetable consumption (oral exposure);
- exposure through inhalation of indoor air, excluding airborne dust particles (inhalation exposure).

For these pathways the following intercompartimental relationships must be derived:

- the relationship between contaminant concentrations in soil and in vegetables;
- the relationship between contaminant concentrations in groundwater or soil, and in indoor air.

Note that no contact medium is concerned with regard to pathway exposure through soil ingestion. In fact, soil material could be regarded as a contact medium in and of itself for this exposure pathway.

Exposure Through Soil Ingestion

Exposure through soil ingestion is an important pathway, especially for immobile contaminants, since exposure through soil ingestion is independent of that part of the contaminant which is in the pore water. Exposure rates are generally higher for children. There is some controversy in regard to the significance of exposure through soil ingestion for adults. Davis and Mirick (2005) concluded that soil ingestion levels in children and adults within the same families were not correlated. Oral exposure through soil ingestion depends on the soil (or soil particle) ingestion rate, the concentration in the soil (or soil particles) and the availability of contaminants in the human body. The pathway exposure through soil ingestion, including the parameter identification, is described in detail in Bierkens et al. (Chapter 6 of this book).

Exposure through soil ingestion by children strongly depends on the children's activities (e.g., Freeman et al. (2005), who investigated the ingestion of pesticides by hand loading on the basis of analyses of the children's activities from videotapes). An important input parameter in the calculation of exposure through soil ingestion is the *relative bioavailability factor* in the human body (e.g., Ruby et al. 1999). This contaminant-specific factor (generally between 0.0 and 1.0) covers the difference between intake (external exposure) and uptake (internal exposure) and reflects the ratio of the bioavailability in the soil matrix as compared to the bioavailability in a reference matrix (e.g., food, or the matrix on which the Toxicological Reference Value is based). This relative bioavailability factor is used to correct for the fact that a contaminant in soil can show a reduced uptake in the body compared to that same contaminant present in food or water. The availability in the human body is dependent on the fraction released from the soil matrix in the stomach during digestion in the gastrointestinal tract (*the accessibility*), the fraction transported across the intestinal epithelium and reaching the portal vein (*absorption*), and the possible metabolism of the contaminant in the intestinal epithelium and/or in the liver. The theory behind bioavailability in the human body and the derivation of the relative bioavailability factor in the human body is described in detail in Cave et al. (Chapter 7 of this book).

Exposure Through Vegetable Consumption

The exposure through vegetable consumption is an important pathway, especially for mobile contaminants. Tsukahara et al. (2003), for example, demonstrated that the consumption of rice was the major source for cadmium exposure in Japan, for 10 out of the 47 provinces (prefectures) where this was investigated. Tao et al. (2004) found PAHs in all eight investigated vegetables, in Tianjin, Northeast China.

Exposure through vegetable consumption is dependent on the representative concentration in vegetables, the total vegetable consumption rates and the fraction of vegetables from the contaminated area to total vegetable consumption. Actually, the relative oral bioavailability in the human body also plays a role for this pathway (e.g., Intawongse and Dean (2006), who investigated the oral bioavailability of cadmium, copper, manganese and zinc through consumption of leaves of lettuce and spinach and the radish and carrot roots), although this process is less important than for exposure through soil ingestion.

It is important to focus on the most relevant concentration in vegetables, that is, the amount of accumulated contaminants in the vegetable at the moment of harvesting and after 'normal kitchen preparation'. Moreover, since the purpose of exposure calculations is to assess the exposure from contaminated sites, only contaminants that originate from soil are of importance. This means that contaminants that are deposited on crops due to rain splash must be included in the exposure calculations, in so far as they are not washed from the vegetables during 'normal kitchen preparation'. However, contaminants from outside the contaminated site that are atmospherically deposited on crops must be excluded from the exposure calculation. This 'outside' atmospheric deposition, however, contributes to the *background exposure* (see Section 5.5.4).

The accumulated concentration in vegetables is often described on the basis of a BioConcentration Factor (BCF), that is, a linear relationship between the concentration in vegetables and in soil. In regard to more sophisticated models for the relationship between contaminant concentrations in soil and in vegetables, it is necessary to distinguish between metals and organic contaminants. The reason for this is that the behaviour in soil and the uptake of these two types of contaminants is fundamentally different. Therefore, different methodologies are used to calculate the concentrations in vegetables for metals and organic contaminants. For metals, more sophisticated models enable the calculation of their concentration in vegetables on the basis of total soil concentration and soil properties (e.g., Krauss et al. 2002). The concentration of organic contaminants is best based on crop type-specific uptake models that follow the Trapp and Matthiess (1995) concept. An example of this model is given in Kulhánek et al. (2005) who calculated the PAH concentrations of leafy vegetables, fruits, root vegetables and potatoes.

Oral exposure through vegetable consumption, including parameter identification, is described in detail in Elert et al. (Chapter 11 of this book). The calculation of the metal and organic contaminant concentration in vegetables, including input parameter identification, is described in detail in McLaughlin et al. (Chapter 8 of this book) and Trapp and Legind (Chapter 9 of this book), respectively.

Exposure Through Indoor Air Inhalation

Exposure through indoor air inhalation is the most important pathway for volatile contaminants. This exposure is dependent on the representative concentration in indoor air and on human characteristics, such as inhalation rate. The representative concentration in indoor air is dependent on advective and diffusive transport of contaminants in pore water and soil air, from the groundwater or the soil into a building and, hence, on the soil properties. It also depends on the building characteristics, such as the possibility for intrusion of contaminated air through holes and cracks, the dimensions of the building and the ventilation characteristics of the building. As in the case of the representative vegetable concentration for the calculation of exposure through vegetable consumption, it is important to focus on the most relevant indoor air concentration for the calculation of exposure through indoor air inhalation. Indoor air concentrations often are characterised by a large variation in time and space (height). Therefore, the calculation of indoor air concentrations typically has a relatively limited reliability. Analogous to the calculation of the relevant concentration in vegetables, for the calculation of the relevant concentration in indoor air, only those contaminants that originate from the soil are of importance. Contaminants that originate from indoor sources (e.g., formaldehyde from furniture, benzene from a garage), however, could contribute to the *background exposure* (see Section 5.5.4). Several models exist to estimate the human health risk due to vapour intrusion, for example, in Australia for the derivation of the Australian Health-based investigation levels (HILs) (Turczynowicz and Robinson 2007). The inhalation exposure through indoor air inhalation, including the parameter identification, is described in detail in McAlary et al. (Chapter 10 of this book).

Insight into Major Exposure Pathways

For the purpose of better understanding the variation in calculated human exposures due to soil contamination, Swartjes (2009) compared the variation in calculated exposures with the variation in calculated concentrations in contact media and in the soil compartments, along with the variation in the input parameters. Serving as the basis for this study were calculations using seven European exposure models for 20 different exposure scenarios. This led to the conclusion that most of the variation in *Exposure through soil ingestion* could be explained by differences in the input parameter *average daily soil intake*. When model-specific input parameters were used, the variation in *Exposure through crop consumption* could be explained by differences in the product of *total consumption rate* and *fraction of total consumption rate that is home-grown*. When standardized input parameters were used, this variation was comparable to the variation in *Concentration in root vegetables* and in *Concentration in leafy vegetables.* The variation in *Exposure through indoor air inhalation* was comparable to the variation in *Concentration in indoor air*. This suggests that the parameters that control the variation in *Concentration in the indoor air*, that is, *surface* and *volume* of the building and, to a lesser extent, *ventilation frequency* of the building, also control the variation in *Exposure through indoor air inhalation.*

5.3.3.5 Other Exposure Pathways

Other exposure pathways that are often included in existing exposure models are (Carlon and Swartjes 2007):

Oral exposure:

- exposure through ingestion of deposited house dust (which partly originates from soil);
- exposure through consumption of drinking water;
- exposure through consumption of animal products, that is, meat, milk and eggs.

Inhalation exposure:

- exposure through inhalation of outdoor air (excluding airborne dust particles);
- exposure through inhalation of indoor airborne dust particles;
- exposure through inhalation of outdoor airborne dust particles;
- exposure through inhalation of vapours during showering.

Dermal exposure:

- exposure through dermal uptake via contact of the skin with soil and dust;
- exposure through dermal uptake via contact of the skin with water during bathing or showering.

Note that no contact medium is concerned with regard to the pathway exposure via dermal uptake via contact of the skin with soil. Just as for the soil ingestion pathway, soil material could be regarded as a contact medium in and of itself.

For some immobile contaminants, exposure through ingestion of deposited house dust can be an important pathway, for example, Lanphear et al. (2003), who showed that exposure through ingestion of settled house dust is a principal source of excess lead among children. For the calculation of the relevant concentration in dust particles, analogous to the calculation of the relevant concentration in vegetables and indoor air, only contaminants that originate from the soil are of importance. Outdoor (airborne) and indoor (deposited and airborne) dust consists partly of local soil material and partly of other materials from other places. The concentration in the dust often is enriched compared to the concentration in the original soil. Again, analogous to exposure through vegetable consumption and indoor air inhalation, contaminants in dust particles originating from indoor sources or places outside the contaminated site contribute to the background exposure.

For the pathways mentioned above the intercompartmental relationships between the contaminant concentrations in soil and in the following contact media must be derived:

- deposited dust particles indoors;
- groundwater, used as drinking water;
- drinking water, after permeation through water pipes;

- meat, milk and eggs;
- outdoor air;
- airborne dust particles indoors;
- airborne dust particles outdoors.

Oral exposure through deposited house dust, as well as input parameter identification, is described in detail in Bierkens et al. (Chapter 6 of this book). The other exposure pathways and intercompartimental relationships mentioned in this section, as well as the input parameter identification, are described in detail in Elert et al. (Chapter 11 of this book).

In general, dermal uptake of contaminants from soil contributes little to total exposure. Contaminants with a relatively high potential for dermal transfer (several PAHs, DDT) can result, however, in significant contributions to exposure through dermal uptake via contact of the skin with water during bathing or showering. In absolute terms, inhalation of dust particles is of minor importance. This route should not be neglected, however, as toxicity by the inhalation route can be significantly different from toxicity by the oral route.

5.3.3.6 Overview Exposure Pathways

In Table 5.1 an overview is given of possible exposure pathways for contaminated sites, including the relevant contact media for each pathway. The pathways are subdivided into oral, inhalation and dermal exposure pathways, and indoor and outdoor exposures.

5.3.3.7 Exposure Scenarios

Any exposure calculation must be based on an *exposure scenario*. Such a scenario describes the site, the soil and, last but not least, human behaviour at the site. With the purpose of promoting understanding of exposure, an exposure scenario could be visually presented by a Conceptual Model.

Two types of exposure scenarios are recognised, that is, *actual exposure scenarios* for site-specific Risk Assessment and *potential exposure scenarios* for generic Risk Assessment (mostly the derivation of Soil Quality Standards).

The basic idea behind *actual* exposure scenarios is simple: mimic the characteristics of the site, the soil under investigation and human behaviour on that site, as well as possible. However, since the outcomes of the Risk Assessment usually need to represent the risks over longer time periods, assumptions must be made for factors that change over time, mainly in regard to the layout of the site and human behaviour characteristics. An example of a transitory factor that relates to human behaviour is the fraction of home-grown vegetables to total vegetable consumption on the site. It is not appropriate to focus on a present situation where no vegetables are grown, since a subsequent resident might be a fanatical gardener, who supplies a substantial part of the total vegetable consumption from the same garden. In this example, it makes sense to define a ‘representative contribution of home-grown vegetables to

Table 5.1 Overview of possible exposure pathways, subdivided into oral, inhalation and dermal exposure pathways, and indoor and outdoor exposure

	Indoor exposure		*Outdoor exposure*	
	Pathway	*Contact medium*	*Pathway*	*Contact medium*
Oral	Dust ingestion	Dust	Soil ingestion	(Soil)
			Vegetable consumption	Vegetables
	Drinking water consumption	Groundwater/ drinking water		
	Meat consumption	Meat		
	Milk consumption	Milk		
	Egg consumption	Eggs		
Inhalation	Air inhalation	Indoor air	Air inhalation	Outdoor air
	Airborne dust inhalation	Indoor airborne dust	Airborne dust inhalation	Outdoor airborne dust
	Water vapours inhalation during showering	Drinking water		
Dermal	Dust contact	Dust	Soil contact	(Soil)
	Water contact during bathing	Drinking water		

total vegetable consumption', for example, 'it must be possible to consume at least 10% of total consumption from one's own garden'. Obviously, this boundary condition is based on a political decision and is very much dependent on the situation of the contaminated site (this differs, for example, for sites in rural areas and in cities). Since site characteristics generally are more stable over time, these can to a large extent be related to the present situation. However, in case the pH, for instance, is strongly influenced by gardening practices, it might be important to include a certain degree of conservatism by adopting a less favourable pH and associated higher plant uptake rates in the Risk Assessment, as in the present situation.

In fact, changes in the contaminant concentrations due to migration and degradation also must be considered. Remarkably, this is not often done in Human Health Risk Assessments. In many Risk Assessments the exposure calculation is based on the present or historical concentration in soil, even when the purpose of the Risk Assessment relates to the human health risks over longer time periods. Since the concentrations in soil generally decrease in time due to leaching and degradation, this can be considered as a worst-case situation. This is not always true for groundwater, however, since leaching from soil could increase contaminant concentrations and contaminant migration could lead to increased concentrations at a distance some ways away from the site. Furthermore, degradation processes may actually worsen conditions (e.g., when TCE is broken down into the more toxic vinyl chloride).

Many of the Risk Assessment factors also must be adapted when determining exposure scenarios in case the Risk Assessment relates to a different future land use or a different layout of the site under the same land use. This situation frequently occurs, especially in densely populated areas such as Northwest Europe, since land use transitions are common and are often preceded by a site investigation and, hence, a Human Health Risk Assessment.

Exposure scenarios for *potential* Human Health Risk Assessment must include the conditions that fit the purpose of the corresponding Soil Quality Standards. In case these Soil Quality Standards, for example, separate the 'no problem' contaminated sites from the sites that may cause an unacceptable human health risk (i.e., a first tier assessment), the exposure scenario must be conservative. It must include, for example, high contributions of the fraction of home-grown vegetables to total vegetable consumption and relate to a vulnerable soil, that is, a soil that shows minimum adsorption (e.g., a sandy soil, poor in organic matter) and, hence, one in which a large contaminant pool is available in the pore water for uptake in vegetables. For Soil Quality Standards that relate to a specific land use, obviously the exposure scenario must apply to that land use. The definition of potential exposure scenarios is a task that involves both decision-makers and scientists. Many of the choices in the exposure scenarios for *potential* Human Health Risk Assessment relate to the political formulation of activities that humans must be able to do at a site without experiencing unacceptable human health effects.

5.3.3.8 Input Parameters

Many input parameters are needed for the performance of a Human Health Risk Assessment. These input parameters can be subdivided into three categories:

- soil-related input parameters (e.g., soil organic matter content, depth of the groundwater table);
- contaminant-specific input parameters (e.g., vapour pressure, solubility);
- human behaviour-related input parameters.

The latter category, human behaviour-related input parameters, can be subdivided into physical input parameters (e.g., body weight, total skin surface), time-activity related input parameters (daily hours at home, shower duration), and contact rates (e.g., soil ingestion rate, total vegetable consumption rate).

There is a different categorisation between site-related and site-independent parameters. The latter (e.g., body weight and shower duration parameters), are often used as defaults in exposure models. Some parameters need to be determined from site-specific information such as the soil-related input parameters, while others could be general or modified for the site (e.g., home-grown vegetable fraction, soil ingestion rate).

In the phase of input parameter identification, an important question is whether or not data that were established in other countries or regions, often with different climates, can be used effectively. Contaminant-specific parameters are generally independent of the site, region or country, the exceptions being sorption characteristics (which could be expressed as a function of soil properties) and plant uptake factors. Physiological factors such as body weight and skin surface area vary only slightly between countries and general data often can be used in Risk Assessment. Since there is no general guidance for dealing with foreign data, a sensible trade-off must be made between the relevance of the data for the Human Health Risk Assessment (e.g., in regard to climate and geographical characteristics for soil-related input parameters or cultural differences for human behaviour-related input parameters) and the number of data points available. In case of lack of data, it is more attractive to include foreign data, even when they have been measured under different conditions.

Exposure characteristics can be found in several documents, for example:

- Exposure Factors Handbook (US Environmental Protection Agency 1997);
- ECETOC 2001, with focus on UK data;
- Otte et al. (2001);
- Child-specific Exposure Factors Handbook (US Environmental Protection Agency 2008);
- ExpoFacts database (Vuori et al. 2006);
- Van Holderbeke et al. (2008).

When choices are made for input parameters for the model, whether it is at the point of developing the model or during site-specific assessments, it is advisable to report background information for the choice made such as the origin of the data, the range of the data and the representativeness.

5.3.3.9 Reliability

Since the mid-1990s, several software packages have become available which stimulate the wide use of exposure models. As a consequence, human exposure models are in widespread use, both implicitly and explicitly. An example of implicit use is the comparison of measured contaminant concentrations with Soil Quality Standards derived from these exposure models (e.g., Swartjes (1999) for the derivation of Soil Quality Standards in the Netherlands; Provoost et al. (2004a) in Flanders, Belgium; and Environment Agency and DEFRA (2002) in the UK). Explicit use is decision-making based on site-specific exposure calculations. Insight into the reliability of calculated exposure could be most directly addressed by performing a validation study, that is, comparing calculated exposure with measured exposure. However, measuring exposure in human beings is difficult or even precluded for technical and ethical reasons. It is possible to validate the calculation of concentrations in contact media such as vegetables and indoor air. From these validation studies it can be concluded that the reliability of these calculations and, hence, of these exposure models is generally limited, because of uncertainties about model algorithms and input parameters. Reliability decreases if contaminants are more mobile and even more so when contaminants are more volatile. The reason for this is that, among the three major exposure pathways, the calculation of exposure through soil ingestion, the dominant pathway for immobile contaminants, is relatively reliable. The calculation of exposure through vegetable consumption, very important for mobile contaminants, is less reliable. And the calculation of exposure through indoor air inhalation, obviously the dominant pathway for volatile contaminants, is the least reliable (Swartjes 2007).

Because of the lack of exposure validation studies, Swartjes (2002), Arcadis (2004) and Swartjes (2007, 2009) compared the results of different exposure models for a standard data set and assumptions. Although such a study does not provide scientific information about the reliability of human exposure models, it does provide insight into the variation of calculated exposure when different models are used. These studies led to the conclusion that the variation in calculated total exposure (combining all exposure pathways) was large. The variation in exposure increases in the following order: exposure through soil ingestion, exposure through vegetable consumption and exposure through indoor air inhalation.

For the promotion of uniformity, it is recommended that a toolbox be compiled for the calculation of human exposure for general use, including standardized tools and flexible tools, the latter to account for region-specific or country-specific (geographical, ethnological, and cultural) elements and national policy decisions.

5.3.3.10 Measurements in Contact Media

Measuring exposure through Biomonitoring is not routine practice in contaminated sites Risk Assessment due to the constraints discussed in Section 5.3.2. However, well-considered measurements throughout the soil-transfer-exposure chain, which could significantly enhance the accuracy of Human Health Risk Assessments, are often done in higher tier Risk Assessments. In regard to exposure due to soil contamination, relevant contact media can be sampled, such as:

- vegetables (e.g., Chunilall et al. (2006), who measured the concentrations of cadmium, nickel and lead in spinach (*Spinacia oleraceaon*), grown on a bituminous coal mine dump soil in South Africa);
- indoor air (e.g., Edwards et al. (2001), who measured VOC (Volatile Organic Carbons) concentrations in residential indoor, residential outdoor and workplace indoor air in Helsinki, Finland);
- dust (e.g., Wilson et al. (2003), who measured floor dust and hand surface wipes to assess the exposure of preschool children, aged 2–5 years, to a suite of organic pesticides and other persistent organic contaminants in North Carolina, USA);
- animal products (meat, milk and eggs) (e.g., Sedki et al. (2003), who measured cadmium, copper and zinc in muscle, liver and kidney of bovines grazing on a municipal wastewater spreading field in Marrakech, Morocco).

These measurements in contact media generally yield more reliable *actual* concentrations than calculated concentrations.

However, it is a misstatement to say that measured concentrations are per definition better than calculated concentrations. The reason for this is that the *actual* concentration is not always a synonym for *representative* concentration. Since concentrations in contact media are fluctuating in time and space, in many cases an average concentration in time and space (or a statistically appropriate concentration, e.g., a 90 percentile) is relevant for Human Health Risk Assessment purposes, while a measurement only relates to a specific place and moment in time. Calculations, although less reliable in estimating actual concentrations in contact media at a specific place, generally focus on an average concentration in time and space.

Moreover, including the variation of the measured concentration in contact media in time and space in Risk Assessment requires special skills of the Risk assessor. Relevant issues include:

- The type and number of vegetables that must be sampled in a vegetable garden. In many small vegetable gardens, however, see Fig. 5.3 as an example, the choice of type and number of vegetables is limited. In this specific case the fast growing crops such as spinach and endive that have a relatively high uptake are not grown on the site.
- The frequency and seasonal planning of indoor air measurements.

By analogy with the position of measurements in body fluids or body tissue, a 'smart' combination of calculating and measuring is the most optimal procedure in

Fig. 5.3 A small vegetable garden in IJmuiden, the Netherlands, as an example of a site which offers the possibility for measuring the concentration in contact media (this is, the edible parts of vegetables), but offers a limited choice of type and number of vegetables for sampling (photo: F. Swartjes)

Human Health Risk Assessment, with a more prominent position for measurements in higher tier Human Health Risk Assessments.

5.3.3.11 Good Exposure Assessment Practice

Several studies are dedicated to the performance of a decent Exposure Assessment such as the *Good Exposure Assessment practice* (e.g., Hawkins et al. 1992; US Environmental Protection Agency 1992; WHO 2005). Like most Risk Assessment elements, good Exposure Assessment practice underwent an evolution from the late

Table 5.2 Ten steps involved in good Exposure Assessment practice (partly based on ECETOC 2001)

	Steps involved	Persons involved
1	Draw-up of a project plan, including a description of the following 9 steps	Project manager, primarily; secondary regulator
2	Determination of the purpose of the Exposure Assessment	Project team
3	Organisation of personnel, funds and (when appropriate) facilities	Project manager
4	Identification of the stakeholders	Project team
5	Identification of political boundary conditions	Project manager and decision-maker
6	Selection of risk tools, including: • special attention to the balance between calculations and measurements (in body fluids and tissues); • a choice for deterministic or probabilistic Risk Assessment; • information about the most important input parameters (including a sensitivity/uncertainty analysis)	Risk Assessment experts
7	Description of the risk tools, including • information about the (validation) status of models, protocols, equations, lists, graphs, etc.; • the procedures for the determination of input parameters (including expert judgement processes)	Risk Assessment experts
8	Description of the outcomes, including: • conclusions of the technical (qualitative) part of the Exposure Assessment • uncertainties involved; • conclusions on the (political) consequences	 • Risk Assessment experts; project manager • Risk Assessment experts; project manager • Risk Assessment experts; project manager; regulator
9	Distribution of and discussion concerning the draft report	
10	Publication of the final report	

1980s until the present time. The major developments since the early Exposure Assessments in risk-based evaluation and management of contaminated sites are the use of tiered approaches, the wider possibilities for using measurements in contact media and, as for any other element of Risk Assessment, the importance of transparency and communication with the stakeholders. Ten steps involved in good Exposure Assessment have been listed in Table 5.2 (partly based on ECETOC 2001). To follow this procedure, it is assumed that a responsible regulator and a Risk Assessment project manager have been identified.

Human exposure modelling would benefit from a higher international consistency in the scientific basis of human exposure models (Swartjes et al. 2009; WHO 2005).

5.4 Hazard Assessment

5.4.1 Contaminants in the Human Body

The hazard of a contaminant is its inherent potential to cause adverse effects when humans are exposed to that contaminant at any level. These adverse effects depend on the nature of the contaminant, the degree of exposure, and the performance of the human body. The *Hazard Assessment* aims to determine the possible adverse effects in the human body. This process includes *hazard identification* and *hazard characterisation* (International Programme on Chemical Safety 2004).

Contaminants may result in *local effects*, related to effects on specific organs at the place of contact or intake, or in *systemic effects*, related to effects in the whole body after systemic circulation and, hence, absorption and distribution in the body. The most common organs or systems that are affected are the lungs, skin, gut, liver (by hepatotoxins), kidneys (by nephrotoxins or renal toxins), nervous system (by neurotoxins), blood, cardiovascular system, immune system (by immune toxins), and the reproductive system. Contaminants that cause cancer are called *carcinogens*. Although there have been major developments in understanding the way cancer cells are formed, the role of contaminants in regard to the development of cancer is still unclear. *Mutagenic* contaminants could be affecting DNA. By changing the DNA in the cell cores, they may be causing cancer or result in, for example, miscarriages as gross chromosomal changes. A common mutagenic contaminant frequently found in soil and groundwater is benzene. *Teratogens* are contaminants that cause birth defects. A common teratogenic contaminant frequently found in soil is lead.

The key word in the Hazard Assessment is *toxicity*, that is, the degree to which a contaminant is able to cause adverse effects. The toxicity of contaminants mainly depends on the reactivity of contaminant molecules in the human body, that is, the capability of forming an association of the molecules with the body's own molecules (a receptor), in this way obstructing vital functions. Therefore, toxicity may be higher for contaminants that mimic the body's own molecules such as a neurotransmitter or a hormone. The toxicity also depends on the role that the body's own molecules play in the vital life processes such as cell replication, embryonic development and the functioning of the endocrine system.

The human body has a whole scale of defence mechanisms for eliminating adverse effects of contaminants that intrude on the human body. First, there are several physical barriers such as the skin, and internal (lipid) obstacles such as the intestine membrane and cell membranes. Second, the body has a number of defensive molecules such as enzymes and vitamins that neutralise toxic contaminants in the body. The next step in the defence system is formed by the specific body organs that are specialised in removing contaminants such as the liver and the kidneys. When the defence system fails and adverse effects result, several repair functions in the human body such as the DNA-repair enzymes, are able to repair primary damage. In spite of all these mechanisms, a part of the intruding contaminants could reach the body's own molecules and cause adverse effects. This could happen, for example, when a surplus of contaminants from a contaminated site enters the body.

Several contaminants that cause severe health damage at high doses may be innocuous or even essential at low doses.

If, and to what extent, primary damage occurs and repair mechanisms are effective is dependent on the genetic talent, life style, dietary habits, age and other factors that determine the physical condition of a human being.

The first step in the Hazard Assessment process is *hazard identification* which defines the type and nature of the adverse effects of the contaminant considered. In the second step, *hazard characterisation*, these adverse effects are quantified and this process ideally results in a dose-response assessment. These dose-response assessments form the basis for a derivation of Critical Exposure values, which are used in Human Health Risk Assessment for risk appraisal.

5.4.2 Threshold and Non-Threshold Effects

As was explained in Section 5.4.1, the exposure rate also is a crucial factor in regard to adverse effects in the human body. In toxicology, an important parameter is the *Toxicological Reference Value*, which represents a benchmark for exposure (*Critical Exposure*) or for a concentration (*Critical concentration*). When this value is not exceeded, it is assumed that there are no unacceptable health risks. When, on the other hand, this value is exceeded, this does not automatically imply that health effects will occur. Exceeding the Toxicological Reference Value generally means that there is *a possibility* for health effects (an unacceptable human health risk). Usually, a further investigation is needed to gain insight into the actual human health risks.

Toxicological Reference Values are generally derived for the oral and the inhalation routes, separately. In case of systemic effects, one Critical Exposure value can be used for both routes, independently from the exposure route, after correction for differences in bioavailability per route. Toxicological Reference Values for long-term systemic effects through dermal exposure are generally not available, but on occasion can be derived from oral Toxicological Reference Values.

The derivation of Toxicological Reference Values depends on the type of associated effects, that is, *threshold* or *non-threshold effects*. Generally, *non-carcinogenic contaminants* are considered as contaminants with a threshold effect. For this type of contaminants, a Critical Exposure value can be derived as a threshold for which adverse effects on humans are unlikely to occur when this Critical Exposure is not exceeded. For *carcinogenic contaminants* the threshold approach is not always applicable. For the category of *non-genotoxic carcinogenic* contaminants, a threshold Critical Exposure may apply. On the contrary, *genotoxic carcinogenic* contaminants, that is, contaminants that directly cause mutations in DNA-material which could result in cancer, are generally considered as non-threshold contaminants. A safe threshold is assumed not to exist for most genotoxic carcinogenic contaminants. In fact, exposure to the smallest possible amount of these contaminants (i.e., one single molecule) increases the risk for cancer.

It must be noted, however, that not all non-carcinogenic effects can be related to a threshold.

The same contaminants can show both threshold and non-threshold effects and the type of effect can differ by route of exposure (oral, inhalation, or dermal). However, if a contaminant shows non-threshold carcinogenic effects, these effects will generally be the most critical effects. In that case, the contaminant is treated as a non-threshold carcinogen.

Traditionally, there has been a lot of discussion about Toxicological Reference Values. The critical effect of low-level lead exposure in children, as one example, is on the central nervous system and relates to cognitive impairment. For years, a threshold of 10 μg lead/dL blood has been considered as safe. Recent research has demonstrated that this threshold should be reviewed and that the absence of a threshold could not be excluded (Chiodo et al. 2007; Rossi 2008). Arsenic, as another example, is a genotoxic carcinogen by the oral route, but discussion upon the presence or absence of a threshold is still ongoing (ATSDR 2007).

Additional details on threshold and non-threshold effects are described in Langley (Chapter 12 of this book).

5.4.3 Toxicological Reference Value for Threshold Contaminants

5.4.3.1 Principles

For threshold effects, a rather universal approach is followed. A level, below which it is believed that no adverse effects will occur, is derived from the *No-observed-adverse-effect-level* (NOAEL) and assessment factors. Alternatively, exposure is compared directly to the NOAEL and a margin-of-exposure or margin-of-safety is calculated, although this method is less useful for the derivation of Soil Quality Standards. Another approach, using the whole dose-response curve, is the benchmark approach. In this approach, the threshold level is derived by applying a statistical regression on the dose-response curve (International Programme on Chemical Safety 1999).

The theoretical options for endpoints in Human Health Risk Assessment are almost inexhaustible, for instance, enzymatic activity, membrane potential, secretion of a hormone, heart rate, or muscle contraction. Human data on reproduction effects, neurological effects, organ toxicity, mutagenesis and carcinogenic effects are available for several contaminants. However, these effect data are often characterized by a poor experimental quality, in some cases, because the conditions of exposure are not always known. Therefore, most Critical Exposure values are derived from animal experiments. Examples of endpoints in animal experiments are alteration of morphology, growth, mass, or life span.

For the purpose of deriving a limit value for exposure (a Critical Exposure value as Toxicological Reference Value), *dose-response relationships* are derived for specific organisms. Such a dose-response curve is derived by exposing the organism to a contaminant at a gradient of concentrations or exposures, often equally spaced on a logarithmic scale (e.g., a concentration of 0.3, 1, 3, 10, 30, 100 mg/kg$_{dw}$). The dose-response relationship often shows an S-curve. In Fig. 5.4 a hypothetical typical S-shaped dose-response relationship is presented. Essential factors in a dose-response curve are the no-effect range, the maximum effect range and, in between, the range of increasing effect with increasing dose. Occasionally, dose-response relationships are U-shaped, such as in case of *hormesis*, for example (*hormesis* is Greek for 'stimulation', i.e., the biological effect at which a contaminant is harmful at higher doses and low dose, but less harmful at intermediate dose; Calabrese and Baldwin (2001)).

The most common effect measurement in animal experiments is the *No-observed-adverse-effect-level* (NOAEL), in other words, the highest experimental dose or concentration at which no adverse effects are shown. In fact the NOAEL is the best available alternative for the actual effect measurement, that is, the *No-adverse-effect-level* (NAEL). When no NOAEL is available, a *Lowest-observed-adverse-effect-level* (LOAEL), in other words, the lowest experimental dose in which the adverse effects are shown, could be used as representative of the NAEL. Since the NAEL is equal to or higher than the NOAEL and always lower

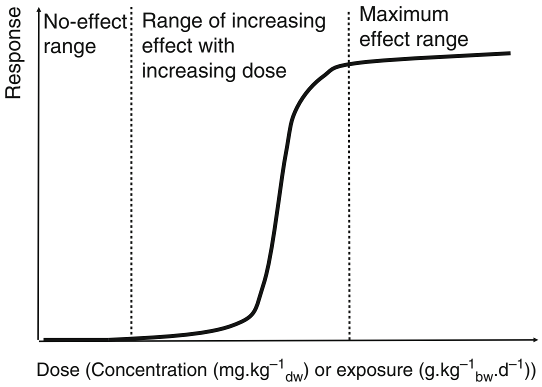


Fig. 5.4 A typical dose-response relationship (S-curve; *solid line*)

than the LOAEL, the NOAEL is a conservative estimate for the NAEL, while the LOAEL is a non-conservative estimate for the NAEL.

In general, there are two options for using experimental animal data as a benchmark in Human Health Risk Assessment. The first is comparing actual human exposures to the NAEL, NOAEL, or the LOAEL from animal studies. Depending on the appropriate margin between exposure and effect dose (MOE: margin of exposure), the contamination at the site will receive a risk classification. However, one should make an a priori decision about that value of the MOE which would represent an acceptable health risk for the contaminant considered, as in Yoshida et al. (2000) who evaluated the human health risks of different population groups exposed to dioxins, in Japan. This decision will relate to the extrapolation from animals to humans, the variability within the human population and the quality of the database. A priori, a decision on the MOE is needed not only for its use in soil quality assessment, but also for consistent Risk Assessment practices. When the actual exposure exceeds these effects measurements, the risk associated with the contaminated sites usually is labelled as 'unacceptable human health risk'. The European Food Safety Agency recommends the use of a *margin of exposure (MOE) approach* for non-genotoxic carcinogens (EFSA 2005).

The second option is more practical and relates to the transformation of the NAEL, NOAEL, or the LOAEL into effects measurements applicable to humans through the use of *assessment factors*. In a more advanced way, the full dose-response data are used to derive a benchmark dose (BMD). The BMD corresponds to a certain level of effect (e.g., 10%). The 95% lower confidence interval (BMDL) is then used as an NOAEL (International Programme on Chemical Safety 2010).

Different names are used internationally for the resulting Toxicological Reference Value, for example, Benchmark exposure, Tolerable Daily Intake (TDI), Acceptable Daily Intake (ADI), Reference dose (RfD), Toxicity reference value (TRV), Reference concentration (RfC), TCA (Tolerable Concentration in Air) and Predicted No Adverse Effect Level (PNEL).

5.4.3.2 Assessment Factors

Assessment factors are used to make effect data on test animals applicable to effect data for human beings. These assessment factors are often assumed to include a whole scale of phenomena such as conversion, extrapolation, adjustment and uncertainty (Vermeire et al. 1999). Use of the overall assessment factor is aimed at converting a threshold dose measured in experimental animals to a safe dose in the human population, including sensitive individuals. Conversion, for example, relates to interspecies differences, that is, differences between the test animals and human beings such as differences in body size and weight, physiology (e.g., related to the bioavailability in the body of organisms and humans) and susceptibility. Extrapolation covers the variation in exposure time between the test animals (often subchronic) and human beings (e.g., lifelong), the differences in exposure conditions between the test animals and human beings, the difference between a high dose and a much lower dose, and the exposure pattern (e.g., continuous exposure of the test organisms versus irregular exposure of human beings in real cases). Extrapolation also relates to intraspecies variation, that is, the differences in sensitivity between human individuals. Adjustment factors could cover, for example, the influence of exposure to more than one contaminant and possible synergistic effects, which causes the overall effect to be stronger than the sum of the separate effects of each contaminant. Uncertainty relates to limitations in the quality of the experimental data set such as the number of tests performed, effect measurement errors, or the use of a LOAEL instead of a NAEL. There is a lot of confusion about terminology in regard to these factors. Extrapolation factors as defined above are often used from a wider angle, that is, as the term assessment factor above.

Another option is to include political and socio-economic factors in the derivation of an overall assessment factor, as is done in ECETOC (1995).

Over the last few decades, there has been a lot of debate about the use of appropriate assessment factors. Many researchers have criticized the poor scientific foundation and the conservative nature of assessment factors (e.g., Slob 1999), while others referred to the lack of uniform terminology and of their use.

In the early days of Human Health Risk Assessment, when the focus was mainly on exposure through food and drugs use, an overall assessment factor of 100 was often used such as by the U.S. Food and Drug Administration (Lehman and Fitzhugh 1954) and the Joint FAO/WHO Expert Committee on Food Additives (JECFA 1987). The factor of 100 was a quantitative representation of a qualitative analysis of differences between test animals and humans. Since then, additional insight has been gained into the true value of assessment factors. The Dutch Health Council, for example, has tried to apply differences in body weight, represented by differences in caloric demand.

The default assessment factor of 100 can be subdivided into two assessment factors of 10, one for interspecies extrapolation and one for intraspecies variation. Each of those can be subdivided into two factors, accounting for toxicodynamic and toxicokinetic aspects. Using the chemical-specific information on toxidynamics and toxicokinetics, modifications in the default sub-factors can be considered (ECHA 2008; International Programme on Chemical Safety 2005).

Finally, the Critical Exposure value is calculated as follows:

$$\text{critical exposure value} = \frac{\text{POD}}{\text{AF}} \tag{5.5}$$

where POD is Point of Departure (e.g., the NOAEL, LOAEL or BMDL) and AF is the assessment factor.

A more sophisticated approach towards assessment factors is the so-called *Benchmark dose concept* in which the separate assessment factors (and the effect levels such as the NOAEL or LOAEL) are expressed as a probability density function and the overall assessment factor is derived by Monte Carlo techniques (e.g., Filipsson et al. 2003; Slob and Pieters 1998).

The derivation of Toxicological Reference Values for threshold contaminants are described in detail in Langley (Chapter 12 of this book).

5.4.4 Toxicological Reference Values for Non-Threshold Contaminants

For genotoxic carcinogens, it is assumed that any interaction of a contaminant with the genetic material in the human body results in a probability of an adverse effect (McMichael and Woodward 1999). Thus, even the lowest exposure would result in a risk for the occurrence of cancer. Although, it is often assumed that a linear relationship between exposure and cancer risk exists over a wide range of exposures, there are indications of deviations from linearity. These are based on the knowledge that cancer occurs through a multi-step process and that DNA repair mechanisms are able to cope with low levels of DNA damage (EFSA 2005).

The approach for the derivation of non-threshold effects is less uniform than for threshold effects (International Programme on Chemical Safety 1999). However, Critical Exposure values for non-threshold carcinogens are mostly derived from animal experiments in which high doses are applied to identify a statistically significant tumour incidence. The dose-response at these high doses needs to be extrapolated to a risk at the lower doses occurring during environmental exposures of humans. This extrapolation often spans several orders of magnitude.

In several guidance documents, Critical Exposure values for non-threshold carcinogens are derived from low-exposure extrapolation, using various extrapolation models. At low exposures, a linear increase of risk with dose is assumed. Therefore, the cancer risk is derived from the *Unit risk* value (i.e., the excess cancer risk per unit of exposure or concentration ($[\mu g \cdot kg_{bw}{}^{-1} \cdot d^{-1}]^{-1}$ or $[mg/L]^{-1}$), aka: *Slope factor*. Unit risk values are generally derived for lifetime exposure. For example, a unit risk of 2.10^{-3} ($[\mu g \cdot kg_{bw}{}^{-1} \cdot d^{-1}]^{-1}$ or $[mg/L]^{-1}$) for a certain contaminant means a risk of 2 excess cancer incidences per 1000 persons with a lifetime exposure to 1 $\mu g/kg_{bw}{}^{-1} \cdot d^{-1}$ or at a concentration of 1 mg/L. This Unit risk can be converted to a Critical Exposure value (Reference dose) once a decision is taken on the acceptable excess cancer risk, as follows:

$$\text{Critical Exposure value}_{\text{genotoxic carcinogens}} = \text{excess cancer risk/Unit risk} \tag{5.6}$$

The decision on the acceptable excess cancer risk is a policy decision. Values used in Human Health Risk Assessment in regard to contaminated sites range worldwide between 1 in 10^4 and 1 in 10^6. Therefore, it is clear that Critical Exposure values for genotoxic carcinogens can easily vary by two orders of magnitude, depending on national policy.

Several alternative approaches are commonly used. The Contaminated Land Policy in UK, for example, uses the *Index Dose approach* for deriving Health Criteria Values for non-threshold carcinogens (Environment Agency 2008). This Index Dose corresponds to the dose expected to be associated with a minimum excess risk of cancer. It is calculated by applying a safety factor to the $BMDL_{10}$ (or T25) from animal experiments, or, if adequate human data are available, by using the dose corresponding to an excess lifetime cancer risk of 1/100,000. The $BMDL_{10}$ corresponds to the lower 95% confidence interval of the benchmark exposure at a

10% response. The T25 corresponds to the chronic exposure rate, which will result in tumours with 25% of the animals at a specific tissue site, after correction for spontaneous tumour incidence (Dybing et al. 1997).

Another approach which is gaining attention in Human Health Risk Assessment for non-threshold carcinogens is the Margin of Exposure (MOE) approach. In that approach, a reference point from the dose-response curve is taken and the actual human exposure is compared with that reference point. The ratio between the exposure at the reference point and actual human exposure is called the MOE and determines the decision on health risk (EFSA 2005). In regard to Risk Assessment at the EU level, the T25 or the BMDL_{10} are recommended as a reference point (EFSA 2005; ECHA 2008). For its application in Human Health Risk Assessment, guidance is needed on the acceptable value for the MOE. EFSA (2005) considers that a MOE of 10,000 or higher based on a BMDL_{10} from an animal study is of low concern. If a T25 is used, the additional uncertainty should be taken into account. This approach is used in the contaminated sites framework of the UK. The resulting Critical Exposure value is defined as an Index Dose (ID).

Critical Exposure values for non-threshold contaminants will be derived for that route or those routes of exposure causing the non-threshold effect.

The derivation of Toxicological Reference Values for threshold contaminants are described in detail in Langley (Chapter 12 of this book).

5.4.5 Reliability

If the Critical Exposure values are derived from well-documented human data, then the accuracy can be high and uncertainty will mainly arise from uncertainty in the dose-response data itself. In comparison, Critical Exposure values derived from animal studies have an uncertainty resulting from the precision in the toxicological studies and from the assessment factors derived. The precision of the Critical Exposure values is inversely related to the assessment factors applied (Speijers 1999). Since Toxicological Reference Values for humans are generally derived indirectly, that is, from experimental animal data and assessment factors, the accuracy of these values is limited. The uncertainty of Critical Exposure values is also reflected in their definitions. As an example, the definition of Reference Dose is given as: ‘The reference dose (RfD) provides quantitative information for use in Risk Assessments for health effects known or assumed to be produced through a nonlinear (possibly threshold) mode of action’. The RfD (expressed in units of $\mathrm{mg{\cdot}kg_{bw}}^{-1}{\cdot}\mathrm{d}^{-1}$) is defined as an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects over a lifetime.

A critical factor with regard to Toxicological Reference Values, and Human Health Risk Assessment in general, is that less knowledge often results in a more stringent Toxicological Reference Value. Or, as formulated in Lindley (2001), the less epidemiologists know about us as individuals, the more at risk the public becomes. It is often claimed that the Toxicological Reference Values are *too* conservative. Dourson et al. (2001), however, showed that a substantial number of Toxicological Reference Values derived from human data are even lower than the values derived from animal studies.

Toxicological Reference Values that are used worldwide vary strongly, as a result of differences in the interpretation of the toxicological studies and in the derivation of assessment factors. Part of this variation is also due to different policy decisions, for instance, in regard to the acceptable choice of excess lifetime cancer risk for genotoxic carcinogens.

Nevertheless, Toxicological Reference Values such as the TDI or Reference dose are useful as long as their derivation is transparent and the exact meaning of the values is realised. A confidence rating on the Critical Exposure value can improve knowledge about the uncertainty for a certain contaminant, as listed in the USA IRIS database. Since the Toxicological Reference Values generally are conservative, it is unlikely that adverse human health effects will occur when Estimated Exposure does not exceed the Critical Exposure. One mistake that could be made is to conclude that health effects will arise if the Estimated Exposure at a site exceeds the Critical Exposure value by any margin. The correct conclusion is that there is an increased risk of health effects and that this increased risk may be considered unacceptable within the regulatory context.

5.5 Risk Characterisation

Risk Characterisation combines the result of the two preceding steps in the Risk Assessment framework, that is, the Hazard and Exposure Assessments. By comparing Estimated (predicted or observed) Exposure to Critical Exposure, a conclusion with regard to human health risk at a site can be drawn.

5.5.1 Site-Specific Risk Assessment

Site-specific Risk Assessment is appropriate if a specific (potentially) contaminated site must be appraised. For site-specific Risk Assessment purposes, the Estimated Exposure relates to the exposure that takes place at that specific site (site-specific exposure). The combination of this site-specific exposure and Critical Exposure, the Risk Characterisation, results in a risk qualification. This risk qualification is either expressed in terms of ‘good’ or ‘bad’ (when the Critical Exposure is not exceeded, or is exceeded, respectively). From a political perspective the risk qualification usually is expressed in terms like *acceptable* or *unacceptable* human health risk. These terms are commonly derived from a scientific range of acceptable levels, combined with political arguments such as ethics, social impact and costs.

In a more sophisticated way, the result of a site-specific Human Health Risk Assessment is expressed as a *Risk Index*, that is, the quotient between the site-specific exposure and the Critical Exposure. Except for an absolute qualification ('good' or 'bad'; acceptable or unacceptable, when the Risk Index is lower or higher than 1.0, respectively) the Risk Index allows for classification of risk qualifications, although this is a subjective process, into classes such as, for example, a 'very high human health risk', when the Risk Index exceeds a value of 10. Moreover, a Risk Index offers possibilities for the scaling of human health risks, which is useful in terms of priority setting. It should be noted, however, that in regard to classification and ranking of risks, the Risk Index assumes a linear relationship with seriousness of human health effects, whereas dose-response curves generally are not linear. Thus a Risk Index of 10 is generally not 10 times worse than a Risk Index of 1.

5.5.2 Soil Quality Standards

Generally, Soil Quality Standards are derived for generic Risk Assessment purposes, that is, for Risk Assessments not related to a specific site. For the derivation of Soil Quality Standards, the Estimated Exposure relates to the *potential* exposure. Potential exposure is defined as the exposure that would occur under specific, standardized conditions in terms of geographical conditions (e.g., relating to soil type, soil properties, depth of groundwater table) and human behaviour (also exposure characteristics: e.g., relating to residence time, amount of home-grown vegetable consumption). To be able to calculate the potential exposure, a generic *exposure scenario* must be constructed, in which the above-mentioned conditions and characteristics are described and quantified.

In regard to the derivation of Soil Quality Standards, the Risk Characterisation is best explained graphically (see Fig. 5.5). In this graph the relevant *potential total*

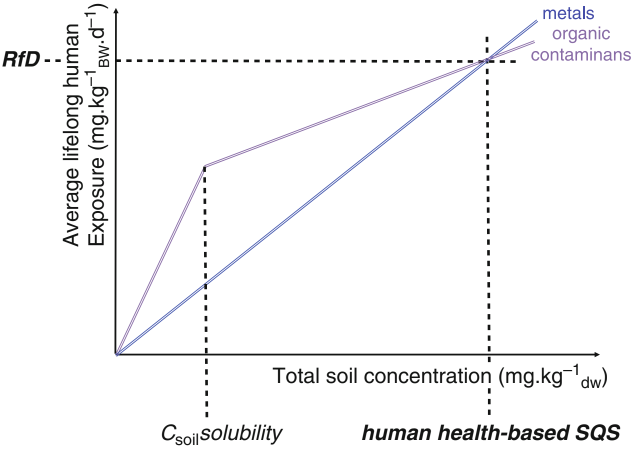


Fig. 5.5 Average life-long human exposure (*potential exposure*) as a function of total soil concentration; the critical exposure (reference dose; RfD); the resulting human health-based soil quality standard (human health-based SQS); C_{soil}*solubility* = concentration in soil at which the water solubility is reached

exposure is given as a function of total soil concentration, both for a metal and an organic contaminant. Moreover, the Critical Exposure for the specific contaminant is given in this graph as a reference Dose (RfD). The human health-based Soil Quality Standard is defined as the total soil concentration that corresponds with a potential exposure equal to the RfD. Note that for metals the relationship between total soil concentration and exposure is assumed to be linear in many existing exposure models. For organic contaminants this relationship often shows a non-linear kink at the soil concentration S_{soil} solubility, at which the water solubility is reached. The reason for this is that at increasing total soil concentrations, the concentration in the pore water remains at its maximum level and exposure via contaminants in the pore water remains at the same level with increasing total soil concentration.

In fact, the procedure described above is a simplification of a more advanced procedure in which oral and inhalation exposures are treated separately, since they generally impact different target organs. To derive the human health-based limit value, or human health-based Soil Quality Standard, the oral and inhalation exposures are related to the oral and inhalation Critical Exposure values. The human health-based Soil Quality Standard is defined as the concentration of a contaminant in the soil for which the sum of the oral and inhalation Risk indices equals 1:

$$(\Sigma \text{ oral exposure}/ \text{RfD}_{\text{oral}}) + (\Sigma \text{ inhalation exposure}/ \text{RfD}_{\text{inhalation}}) = 1 \tag{5.7}$$

where Σ oral exposure is the sum of exposure from all oral exposure pathways [$\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$], RfD_{oral} is the Reference dose for oral exposure [$\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$], Σ inhalation exposure is the sum of exposure from all inhalation exposure pathways [$\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$], and $\text{RfD}_{\text{inhalation}}$ is the Reference dose for inhalation exposure [$\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$].

The Reference dose for inhalation exposure is generally derived from the Reference concentration (RfC), or Tolerable Concentration in Air (TCA).

Strictly speaking, the above equation is only applicable in the case of contaminants where the critical effect is systemic (i.e., affecting the whole body) and is the same for all exposure pathways. If a local effect is considered critical for one of the exposure routes (e.g., exposure through inhalation) or the critical effect differs per route (e.g., carcinogenicity for one of the exposure pathways), then exposures from the oral and inhalation routes should not be combined in the Risk Characterisation.

Dermal exposure requires special attention. Dermal exposure is directly calculated as an absorbed dose, while Critical Exposure values are often not available. Therefore, dermal exposure is generally combined with oral exposure. However, the metabolisation and differences in absorption between the oral and the dermal routes actually should be accounted for to the fullest extent possible. A Critical Exposure value for systemic effects through the dermal exposure routes could be calculated from the oral Critical Exposure value, and the oral bioavailability or absorption factor (US Environmental Protection Agency 2004). Alternatively, the absorbed dermal exposure could be converted into a pseudo-oral exposure by multiplying it with the same oral bioavailability factor and then combining it with the oral exposure pathways. Corrections for bioavailability should only be made, however, if gastro-intestinal absorption is less than 50% (US Environmental Protection Agency 2004).

In spite of the inaccuracy involved with the use of the general equation (5.7), this simplified approach is used in many regulations and in guidance for contaminated sites. Guidance for a more detailed Risk Characterisation is available from other frameworks (e.g., ECHA 2008).

In addition to comparing exposures, concentrations in environmental compartments can be compared with specific limit values. These limit values can be toxicologically based (e.g., limit value of drinking water, or limit value of indoor air), but can also have a regulatory basis (e.g., limit values for food stuffs). Although exceeding these limit values is generally considered unacceptable, it cannot be concluded a priori that there is an unacceptable human health risk involved. Limit values in food stuffs generally combine a toxicological evaluation with the principle of 'as low as possible under good agricultural practices'. For these limit values, concentrations above the limit values will not result in an unacceptable health risk if total exposure is below the Critical Exposure value.

Moreover, Toxicological Reference Values are often based on the assumption that continuous exposure is taking place, while for most scenarios exposure only takes place during a part of the day. On the other hand, politically-based limit values could also reflect feasibility and be higher than what would be considered an acceptable toxicologically based risk limit. In that case, a conflict between the purely toxicological approach for contaminated sites Risk Assessment and politically-based limit values could arise.

5.5.3 Relevant Time Span

After calamities, human beings can be exposed over a short period to often high contaminant levels. Analogously, humans are exposed for short periods when they spend a relatively short time at a contaminated site, for example, when children play every now and then at a contaminated industrial site or when humans recreate for short periods in a contaminated area. In many other cases, humans are exposed for longer periods and on a more continuous basis, from several years up to their whole lifetime, for example, when they live or work on a contaminated site. In other words: exposure ranges from *acute*, via *sub-acute* and *sub-chronic*, to *chronic*. There are no absolute definitions for these adjectives (International Programme on Chemical Safety 2004). However, acute effects often refer to effects that occur within 24 h or within the 14 days after exposure; sub-chronic and chronic exposure in experimental toxicology is often defined as 'exposure up to 90 days' and 'considerably more than 90 days', respectively. The phrases 'acute exposure' and 'chronic exposure' are also often used in the meaning of single exposure and repeated exposure, respectively, but this distinction does not leave much scope for further differentiations in regard to exposure duration.

Obviously, it is essential to relate site-specific Human Health Risk Assessment to actual exposure times. Since the amount of time that humans will spend on a contaminated site is not always known, the definition of the relevant exposure time also is a political issue. It could politically be stated, for example, that in principle it should be possible for humans to work at a contaminated site for a period of 40 years (working lifetime), without experiencing unacceptable adverse effects due to soil contamination.

The political role is even more pronounced for the derivation of Soil Quality Standards. As was described in Section 5.5.2, one or more generic exposure scenario(s) have to be defined as an important basis for the derivation of Soil Quality Standards. In fact, the exposure time, or the duration that humans perform different tasks at a contaminated site, must be part of this exposure scenario. According to the Dutch Soil Protection Act, for example, a political boundary condition for the derivation of Soil Quality Standards (*Intervention Values*, used in the first tier Risk Assessment) is that humans must be able to reside their whole lifetime on a contaminated site. Although this is hardly a realistic condition, certainly in case of any land use other than Residential, it is believed to represent an acceptable worst-case (conservative) criterion to be used in the first tier Risk Assessment. An exception is made for lead for which only the child phase is considered, since human health effects are much more relevant during the child phase.

Time span should not just be considered in the Exposure Assessment; it is also a key component in Risk Characterisation, as Estimated Exposure and Critical Exposure values should be consistent with regard to duration. Exposure Assessments at contaminated sites generally address long-term exposures, thus averaging out any peak exposures occurring at the site. This is generally appropriate for residential or industrial sites, as peaks in exposure are not expected. One exception is pica behaviour, particularly by small children, where deliberate ingestion of large amounts of soil particles could result in peak exposures that go well beyond typical assumptions (e.g., Lemanek et al. (2002), who examined the incidence and relationship of pica symptoms and dysfunctional eating patterns in children and adolescents with sickle cell disease).

Scenarios with intermittent exposure, for example, in contaminated recreational areas where exposures take place only for a few hours per day, for a few days a year, would, however, require consideration of the potential health effect of this short-term exposure.

A third aspect in regard to time span is the toxicologically relevant exposure period in the lifetime of the exposed population. Lead, for example, is well-known to cause neurological effects by children, but can affect the kidneys at higher and longer exposure levels. Moreover, health effects can be related to the average daily dose, but can also be related to the cumulative dose.

Susceptibility can change with age, and much attention is paid to the susceptibility of children or the elderly for exposure to contaminants (International Programme on Chemical Safety 2006; US Environmental Protection Agency 2006). A number of pieces of legislation account for the protection of children by performing the Risk Characterisation of threshold contaminants upon exposure during the first years of age (0–6 years). This practice is applied, for example, in the United Kingdom and in Flanders (Belgium). As children will have a higher exposure per unit of body weight

(Dybing et al. 1997), a Risk Characterisation based on children will result in a more stringent Risk Assessment. In the Netherlands, lifetime average exposure is calculated, except for the assessment of lead exposure, where the child is also considered to be part of a susceptible subpopulation. Grosse et al. (2002), for example, showed that because of falling lead-blood levels, U.S. preschool-aged children in the late 1990 s had IQs that were, on average, 2.2–4.7 points higher than they would have been if they had the blood lead distribution observed among U.S. preschool-aged children in the late 1970s.

For commercial sites, the United Kingdom and Flanders (Belgium) limit exposure duration and averaging time to a high-end estimate of the duration of a professional career. With regard to non-threshold carcinogens, the general approach is that a lifelong average exposure should be calculated, even if the toxicologically relevant exposure duration is shorter than a lifetime.

5.5.4 Background Exposure

Humans are in constant contact with contaminants. They eat and drink contaminant-holding foods. Contaminants can enter the food chain during primary production (e.g., pesticide residues in vegetables from the super market) or are created during the production of food (e.g., Polycyclic Aromatic Hydrocarbons from the grilling of meat). Humans inhale contaminated air or volatile contaminants from glues, paint, petrol, and printed works. They also put contaminant-holding materials on their skin such as cosmetics and lotions.

From the perspective of exposure on contaminated sites, these exposures can be considered as *background exposure*. Falcó et al. (2004), for example, showed a significant contribution of hexachlorobenzene exposure through the consumption of dairy products for the population of Catalonia, Spain. Cetin et al. (2003) demonstrated elevated VOC-concentrations in ambient air, mainly of ethylene dichloride, ethyl alcohol and acetone, from a petrochemical complex and oil refinery in Izmir, Turkey.

It makes an important difference if background exposure is of a voluntary or imposed nature. Voluntary exposure, such as exposure to cadmium through smoking, or exposure to alcohol in drinks, can be controlled by humans. Exposures to soil contaminants or to industrial emissions are typically imposed exposures.

Typically, the amount of background exposure is very contaminant-specific. Moreover, background exposure differs among regions, depending on food pattern, lifestyle, traffic density, and the presence of urban or industrialised areas. Therefore, background exposure varies on a regional scale and is specific for different countries or regions. Several sources give estimates of background exposure for specific regions.

From a *medical viewpoint*, total exposure (i.e., exposure from the contaminated site, and voluntary and imposed background exposure combined) must be regarded in cases of contaminants showing threshold effects. The reason for this is that the human body does not distinguish between exposure from background sources and that from the contaminated site. From this perspective, reduction of exposure from the contaminated site or from background sources is of equal importance. Intervention measures should ideally focus on the most effective way of reducing exposure. Often reducing voluntary background exposure is the most promising and effective.

Whether or not to include background exposure in Risk Assessment and Risk Management is, however, a political issue. From a political perspective, different viewpoints on the role of (imposed) background exposure in Human Health Risk Assessment are possible. One viewpoint follows the medical approach, the most pure approach from a toxicological perspective. This implies that exposure from the contaminated site must be added on to the (imposed) background exposure and the combined exposure may not exceed the specific Critical Exposure. Or, in other words, exposure due to contaminated sites may not exceed a Critical Exposure (aka: Toxicological Reference Value), reduced by the (imposed) background exposure (see Fig. 5.6).

One practical problem is that for some contaminants the (imposed) background exposure fills up a substantial part of the Critical Exposure values and sometimes even exceeds the Critical Exposure values. In that case, the human health risks due to contaminated sites cannot be assessed on the basis of 'the medical approach'. This aspect could result in conflicts when setting Soil Quality Standards. In Flanders, for example, background exposure is taken into consideration (Provoost 2004a, b) and could therefore constitute a substantial part of total exposure. In the UK (Environment Agency 2008) and in Germany (Bachmann et al. 1999), background exposure is limited to a specified proportion of the Critical Exposure value for every contaminant.

Another principle refers to the fact that the policy on contaminated sites can only control exposure due to contaminated sites, and must take the (imposed) background exposure for granted. According to this philosophy, it is politically defensible to only assess the exposure due to contaminated sites. In other words, independent of background exposure and, hence, of the overall effects, the exposure due to contaminated sites may not exceed Critical Exposure. Since humans may not be 'fully protected' according to this philosophy, and it would thus not be correct to state that it is safe to reside on the site, a clear communication to that effect is needed to explain the reasoning behind this political position.

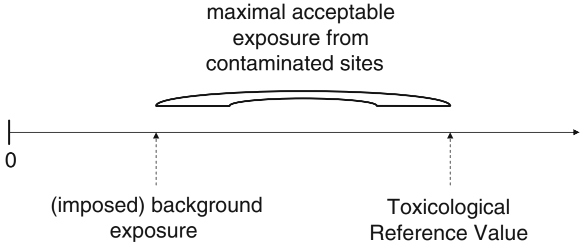


Fig. 5.6 Maximum acceptable exposure from contaminated sites, as critical exposure (aka: toxicological reference value) reduced by the (imposed) background exposure that may not be exceeded

One alternative possibility is to indirectly include background exposure by using an appropriate Critical Exposure, for example, by using additional safety factors (often of 10 or 100). Such a limit value is at the protection level of Negligible Risk, for instance. This risk level is supposed to cover, among other things, the effect of background exposure.

The role of background exposure is different for *non-threshold carcinogens*. For these contaminants, an acceptable risk level is set, which corresponds to the *excess* risk that is accepted for soil contamination. Therefore, background exposure does not need to be incorporated in the Risk Characterisation.

5.5.5 Combined Exposure

In the great majority of contaminated sites show contamination with more than one contaminant in soil and/or groundwater. Sometimes, the same combinations of contaminants are found in soils and groundwater, at other sites an incoherent cocktail of contaminants is present. As a consequence, humans generally are exposed to more than one contaminant at the same time.

Non-carcinogenic contaminants can act independently from each other, or can influence the overall effect due to combined exposure. Although the exposure rate differs for each contaminant, the impact of this combined exposure must be assessed in a Human Health Risk Assessment. Exposure to more than one contaminant can increase the human health risk. For the magnitude of this increment, three different possibilities exist, depending on the composition of the contaminant cocktail. According to the first possibility, contaminants do not influence the potency of the other contaminants. In this case, two types of addition exist, *dose addition* and *response addition*. For contaminants with the same toxicological endpoint (e.g., target organ) that act through a common mode of action, dose addition is appropriate. For these contaminants the doses can be summed up, if necessary after accounting for differences in potency, and then compared to the Critical Exposure. If contaminants have the same endpoint, but act through a different mode of action, response addition applies. In that case, the responses should be added. In practice, this means that the combined effect of contaminants showing dose addition can lead to a negative health risk appraisal, even if the separate exposures do not exceed the Critical Exposure. In contrast, the combined effect of contaminants showing response addition will not exert a negative health risk appraisal when the separate exposures do not exceed the Critical Exposure (Wilkinson et al. 2000). The second possibility is that the exposure to several contaminants enhances the overall effect more than linearly, which are called *synergistic* effects. Third, a less than linear increase in effects is possible, known as *antagonistic* effects. Quantification of synergy and antagonism is difficult and, therefore, seldom done. In practice, linear addition of exposure is often performed for contaminants that have the same mode of action.

As a simplified approach to dose addition, the different Risk indices (ratio of exposure to Critical Exposure value) are added up, while the criterion for ‘possible unacceptable human health risk’ is:

$$(\text{exposure}_1/\ \text{RfD}_1) + (\text{exposure}_2/\ \text{RfD}_2) + \ \ldots\ + (\text{exposure}_n/\ \text{RfD}_n) = 1 \qquad (5.8)$$

where

$\text{exposure}_1, \text{exposure}_2, \ldots, \text{exposure}_n$ is exposure to contaminant 1, 2, . . ., and n [$\text{mg.kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$];

$\text{RfD}_1, \text{RfD}_2, \ldots, \text{RfD}_n$ is Critical Exposure of contaminant 1, 2, . . ., and n [$\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$].

From a quantitative Human Health Risk Assessment perspective, this procedure is not correct, since it falsely assumes that the risk is linearly related to the Risk Index (e.g., that a Risk Index of 0.6 is two times worse than a Risk Index of 0.3). In spite of the lack of scientific foundation, Eq. (5.8) offers a practical way to account for an increased risk when more contaminants are present.

In case of effects due to exposure of contaminants with a different endpoint, two or more different threats to the human health risk must be accounted for. However, since there is no standard procedure to quantify the overall effect of two different human health threats, this is rarely done in Human Health Risk Assessments. Generally, attention is focused on the contaminant that is expected to cause the most serious health effect.

Comparable considerations can be made for non-threshold carcinogens. If they act through the same mode of action on the same endpoint, exposure could be combined. However, even in regard to different targets and modes of action, there may be a potential for combined effects (Environment Agency 2008).

Typically, in practice effects from carcinogenic contaminants are often assumed to be additive, regardless of target organ.

5.6 A Closer Look at Human Health Risk Assessment

5.6.1 Significance of Exceeding Toxicological Reference Values

If exposure due to soil contamination exceeds the Critical Exposure value, the question arises as to what the significance of this exceeding means in terms of health effects. The question is whether adverse health effects will indeed occur for the human beings involved.

Exceeding a Toxicological Reference Value should be interpreted on a case-by-case basis. The conclusion will depend on the toxic mechanism for the critical effect, the toxic endpoint, the margin between effects at chronic and sub-acute/sub-chronic exposure and the potential for different toxic effects at sub-acute/sub-chronic exposure.

Generally, Toxicological Reference Values are set at safe levels that protect the general population, including sensitive individuals. The protective nature of an ADI, for example, predominantly depends on the accuracy of the toxicological studies, as well as on the range and number of doses tested (Speijers 1999). In general, the safety factors used are higher as less information is available. If exposure occurs over a short term, a slight exceedance of a Toxicological Reference Value is not expected to result in a substantial human health risk. This is different for an exceedance of a Toxicological Reference Value over longer time periods, as overall safety is reduced and sensitive population subgroups become less well-protected (ILSI 1999). The dose-effect relationship for threshold contaminants mostly follows an S-shaped curve (Fig. 5.4). Thus, the consequence of exceeding the threshold value will depend on the steepness of the slope of the curve and where on the curve any particular exposure falls. An assessment of the increased risk for long-term exceeding of a long-term Critical Exposure value can be made by comparing NOAEL and LOAEL values or more explicit use of the dose-response curve.

If short-term exceeding of the Critical Exposure values occurs, as can be the case in intermittent exposure scenarios, the Critical Exposure value for long-term exposure can generally not be used to assess human health risks. In cases where a contaminant with high accumulation potential in the human body is concerned (e.g., cadmium or dioxins), and there is a large margin between the chronic NOAEL and the acute toxic dose, the impact of the cumulative dose can be easily calculated and an unacceptable human health risk is not expected as long as the critical cumulative dose is not exceeded. If the half-life of a contaminant is short, however, the risk results from chronic stress and the risk from short-term peak exposures should be assessed using the evaluation of short-term toxicity studies (Walker 1999).

International bodies have published Critical Exposure values for threshold effects for short-term exposure values such as the MRLs (Maximal Risk Levels) from ATSDR (which also published longer term MRLs).

5.6.2 Odour Nuisance and Taste Problems

Contaminants in soil or groundwater may cause *odour nuisance* or *taste problems*, in addition to toxic effects. There is no significant association between odour perceptibility and the potential of the contaminants present to induce health effects (Rosenkranz and Cunninghal 2003). Non-odorous contaminants, however, are more likely to cause adverse human health effects. Therefore, the authors concluded that odour cannot be used as a sentinel for health hazards. Nevertheless, odour and taste problems might be aesthetically unpleasant and annoying. It is a policy decision if odour nuisance and/or taste problems should be used as criteria for contaminated site appraisal or as a criterion in the design of Remediation objectives.

Although human perception strongly differs among individual humans, odour and taste limit values for several contaminants do exist. The corresponding soil or groundwater concentration, however, is for some contaminants very stringent as compared to human health or ecologically based risk limits. Swartjes et al. (2004), for example, derived odour and taste limit values for MTBE in groundwater of 15 and 40 μg/l, respectively, while they derived a toxicologically based limit value in groundwater of 9420 μg/l.

5.6.3 Physiologically-Based PharmacoKinetic Modelling

With the exception of some medical and occupational data, reliable human data on health effects related to specific exposure levels are rarely available. Therefore, as noted previously, potential human health effects are mostly derived from animal experiments. As a consequence, animal effect data have to be extrapolated to project data on humans, using assessment factors to accommodate for factors such as the effects on one population to the other, from short-term high dose to long-term low dose, etc. Due to lack of scientific knowledge, these extrapolation factors are often put at 10, 100 or 1000, generally considered as conservative extrapolation factors (see Section 5.4.3.2 for more details).

In some situations, the prediction of internal exposure or body burden by use of *Physiologically-Based PharmacoKinetic* (PBPK) models can provide additional value. These models focus on a more realistic 'biologically effective dose', in other words, the exposure which casually relates to effects. A PBPK model is a tool to estimate the exposure in organs, body tissues or fluids for which Toxicological Reference Values exist. PBPK models describe the fate and transport of contaminants in the human body, that is, the rate at which contaminants are adsorbed, distributed, metabolised, excreted (renal excretion) and eliminated (hepatic elimination). For this purpose, these models seek to mimic the physiological and biochemical processes of the human body, and include equations on transport in the blood, partitioning into tissues and enzymatic conversion. PBPK modelling also offers the potential to extrapolate outside the range of experimental conditions.

PBPK modelling has been used in medicines, for many decades. Its application in human health contaminated site management, certainly in regard to formal regulations, has so far been limited. The poor availability of suitable pharmacokinetic models for a variety of contaminants hampers their use in contaminated sites Risk Assessment. Moreover, enhanced data requirements limit their use, certainly in lower tier Risk Assessments. Leggett et al. (2003), as an example, developed a PBPK model for the purpose of investigating the fate of caesium in the human body. That model, which was constructed around a detailed blood flow algorithm, includes the calculation of the transfer of caesium from the blood plasma into the tissue and vice versa, and the secretions into the gastrointestinal tract. The British 'Risk Assessment and Toxicology Steering Committee' endeavoured to arrive at a more standardized approach to formal frameworks (Government/Research Councils Initiative on Risk Assessment and Toxicology 1999). Cornelis et al. (2006), as another example, used the Integrated Exposure Uptake Biokinetic (IEUBK) model in order to estimate blood lead concentrations of children exposed to lead in the vicinity of a non-ferrous plant situated in Hoboken, Belgium.

For larger and more polar contaminants, for which body membranes offer more resistance to permeation, more insight in the fate and transport processes in the human body is especially needed. For this purpose, the development of databases for basic morphological and physiological parameters has been encouraged.

5.6.4 Probabilistic Human Health Risk Assessment

Since Human Health Risk Assessment can be characterized by large uncertainties, an attractive alternative to deterministic Human Health Risk Assessment is *probabilistic Human Health Risk Assessment*. The basic principle here is to replace point estimates of input parameters with probability density distributions. Popular distributions include uniform, log normal and triangular distributions. In performing the calculations, called Monte Carlo calculations, values for the input parameters are randomly extracted from these distributions. The result of the Human Health Risk Assessment is a distribution of calculated exposures or Risk indices. Probabilistic approaches are particularly popular for the calculation of exposure variability, for example, for the Risk Characterisation for dibutylphthalate, as in Vermeire et al. (2001).

The big advantage of probabilistic modelling is that it shows the possible range of exposures or Risk indices and thus provides insight into the variability and uncertainty of the estimates.

One disadvantage of probabilistic modelling, however, is that a choice must be made for the percentile of the protected population. Percentiles that are often chosen, in most cases without a clear underpinning, are the 50th (the medium exposure), the 80th, 90th, or 95th percentile. Although it is often claimed that this represents a political decision, scientists must play a role in explaining the exact meaning of and consequences of different choices for a specific percentile.

Ane interesting criticism of probabilistic exposure modelling is that epidemiologists use probability to redistribute human anxiety; Lindley (2001) claimed that, typically, they falsely reassure some and baselessly frighten the rest. This author decided in favour of black and white decisions (good or bad). The only difference between black and white decisions and probabilistic approaches is, according to the author, that the former is decided on the antecedent conditions, while the probabilistic approaches translate the unknown antecedent conditions into potentially confusing information.

One alternative to Monte Carlo techniques is using fuzzy-stochastic modelling approaches. In this approach, fuzzy membership functions are employed to quantify the uncertainties and complexities. An example is given in Chen et al. (2003), who applied this procedure to evaluate the human health risks resulting from subsurface toluene contamination at a petroleum-contaminated groundwater system in western Canada.

5.6.5 Reliability

In Section 1.5.4 a general exposition on reliability is given. In human exposure modelling, the model user has to deal with many input parameters. It typically is not efficient to determine each input parameter with the same degree of precision. In good modelling practice, the model user pays specific attention to the identification of the most important input parameters. The most important parameters are those parameters for which it is generally known that they play a crucial role for the specific Risk Assessment (e.g., a vegetable-soil BioConcentrationFactor, when the Risk Assessment focuses on the human health risk related to vegetable consumption from a contaminated site) or ones that have a large influence on the outcome of the Risk Assessment (e.g., on a Soil Quality Standard or a Risk Index). The first aspect, that is, identifying the most crucial parameters for the specific Risk Assessment, mainly is a matter of experience. An experienced model user knows from the contaminant characteristics which exposure pathways and corresponding input parameters are the most relevant. The second aspect, that is, the assessment of the influence of input parameters on the outcome of the Risk Assessment, can be investigated by means of a *sensitivity and uncertainty analysis*. A sensitivity analysis investigates the impact of a predefined change of an input parameter (e.g., a 10% increase) on the outcome of the Risk Assessment. An uncertainty analysis investigates the impact of an input parameter, given its distribution, on the outcome of the Risk Assessment.

5.6.6 Ethical Issues

5.6.6.1 Human Beings

The most direct way to test the effects of contaminants in regard to human health effects would be to expose humans to contaminants under controlled conditions. This kind of experiment, for example, *human volunteer studies* in which the volunteers are often paid, will also lead to the most accurate Critical Exposure values. However, testing human beings has been recognised as a critical ethical issue for centuries. In the last century, several codes of law were released on this topic, starting with the Nuremberg Code, following the Nuremberg Military Tribunals after World War II. The major question inherent to this debate is: ‘under what circumstances is it morally and ethically acceptable to intentionally expose humans to contaminants, for which it is known that they can experience adverse human health effects?’ There is recognition among the general public that human experimentation is a critical aspect of the development of new drugs and approaches to treat diseases (Roberts 2001). However, there is much more debate on human experimenting in regard to Human Health Risk Assessment related to environmental issues such as contaminated sites (e.g., McConnell 2001). Most certainly, when investigating an effect dose, the experimental human beings must experience adverse effects by definition.

Conditional supporters of human experiments refer to the human health effect balance: by affecting a few, many may benefit. Koren and Seal (2001) noticed that many of the public human health decisions, policies and practices in the USA have benefited from, and have been based on, research wherein human beings are exposed to some degree of risk. In the USA, a document was published with recommendations on circumstances when dosing of humans with contaminants would and would not be acceptable (US Environmental Protection Agency 2000).

Epidemiologists prefer to use exposure situations already existing in society, although the disadvantage of these ‘studies’ is that these exposure conditions are not always favourable and often are even unknown.

5.6.6.2 Animals

The last few decades, there has been an intensive debate on the use of animals, such as mice, rats and pigs, for experiments. It is often stated that animal testing is cruel and leads to the unnecessary suffering and deaths of hundreds of thousands of animals worldwide. The majority of these experiments is performed for medical purposes. But these types of experiments are also performed for environmental quality assessment. Several organisations such as the *European Centre for the Validation of Alternative methods* (*ECVAM*) in the European Union (ECVAM 2009), *People for the Ethical Treatment of Animals* (*PETA*) in the US (PETA 2009) and Animal Aid in the UK (Animal Aid 2009) focus on the fate of experimental animals. Today, it is widely recognized that animal experiments must be prevented as much as possible.

5.6.7 Relationship Scientist and Decision-Makers

It is often said that Risk Assessment is an objective process and that scientists must operate independently of the interests of any stakeholder. To a certain extent this is true; scientific independence is a key to an objective risk qualification. This independent position, however, certainly does not justify a strict ‘no communication policy’. The independent status of scientists will not be affected by adaptation to specific political boundary conditions, as long as it is made transparent what these boundary conditions are. Human Health Risk assessors can do an excellent and objective job when they, for example, commit themselves to the political boundary condition that a Human Health Risk Assessment for an industrial site focuses on ‘average adult workers’ and does not relate to children or other sensitive groups. Again, it is important to identify these boundary conditions and, hence, make the validity range of the conclusions from the Risk Assessment transparent. Therefore, this political boundary condition must clearly be described in the Risk Assessment report. This enables regulators to ensure the safety of these sensitive groups, for example, by fencing off the site using some means of controlled access to protect children in the case of the above-mentioned example.

5.6.8 Site-Specific Risk Assessment

The most elegant way of dealing with site-specific Human Health Risk Assessment is by using a *tiered approach*, that is, a stepwise approach. In such an approach several assessment steps (tiers) are described. In each higher tier, the assessment is based on more site-specific elements, is less conservative, and, hence, is more complex, time-consuming and often more expensive. The philosophy behind this is: 'simple when possible (only the first tier) and more complex when necessary (higher tiers)'. A tiered approach represents an efficient way of Risk assessing without compromising scientific integrity.

The use of a tiered approach for Human Health Risk Assessment is often applied. Today, the application of generic Soil Quality Standards often is the first step in this tiered approach. Further tier assessments can be undertaken to enhance the Exposure Assessment by making use of site-specific measurements (e.g., concentrations in vegetables, indoor air, or dust), Biomonitoring, running more complex models (e.g., site-specific vapour intrusion models) or refining exposure calculations. The latter is possible, for example, by including site-specific exposure information (e.g., time-activity data, local food consumption, oral bioavailability in the human body), or by applying Physiologically-Based PharmacoKinetic models. Furthermore, exposure metrics can be tuned to the toxicological properties of the contaminants (critical time-frame, effects by exposure route).

One example of a tiered approach in Human Health Risk Assessment is shown in Swartjes et al. (2007), for the determination of human health risks due to the consumption of vegetables grown on contaminated sites (Fig. 5.7). In this procedure four steps have been distinguished, that is:

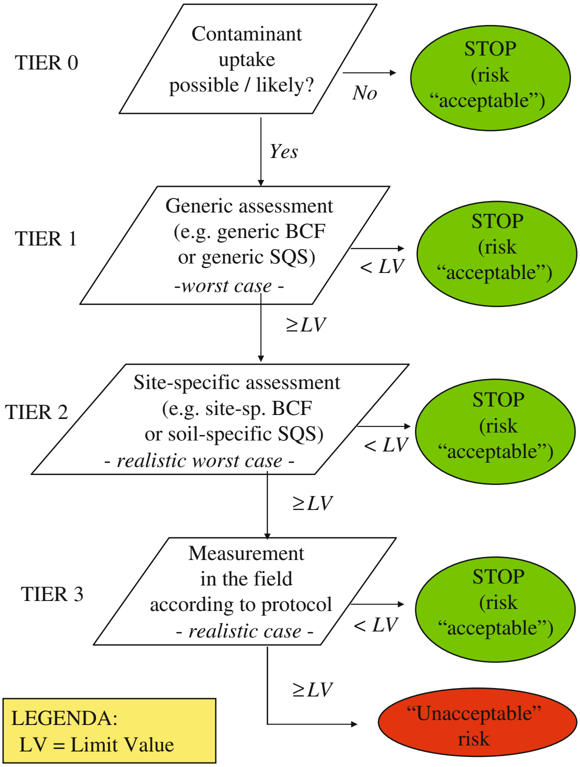


Fig. 5.7 The four step (tiered) procedure for the determination of human health risk due to the consumption of vegetables grown on contaminated sites used in the Netherlands, as an example of a tiered approach in Human Health Risk Assessment (Swartjes et al. 2007)

- testing the probability of consumption from contaminated sites (tier 0; this tier is called ‘Tier 0’ instead of ‘Tier 1’, since no Risk Assessment is performed; the only argument is whether contaminant uptake by plants is possible or likely, for example, given the location of the site);
- comparing measured soil concentrations with conservative ‘critical soil concentrations’ (Tier 1);
- site-specific calculation of the vegetable concentration on the basis of soil properties (Tier 2);
- measurement of the vegetable concentration following a specific protocol (Tier 3).

References

Animal Aid (2009) http://www.animalaid.org.uk/h/n/AA/HOME/. Retrieved 24 Nov 2009

Arcadis (2004) Risk assessment comparison study. Arcadis Geraghty and Miller International, report number 916830024, April 2004

ATSDR (2007) Toxicological profile for arsenic. Agency for toxic substances and disease registry, United States. Available via http://www.atsdr.cdc.gov/toxprofiles/tp2.html#bookmark07. Cited 15 Dec 2008

Bachmann G, Oltmanns J, Konietzka R, Schneider K (1999) Calculation of Screening values for the assessment of historical soil pollution (in German). Umweltbundesamt, Erich Schmidt, Berlin

Brand E, Otte PF, Lijzen JPA (2007) CSOIL 2007: an exposure model for human risk assessment of soil contamination. RIVM report 711701054 RIVM, Bilthoven, the Netherlands

Budd P, Montgomery J, Evans J, Trickett M (2004) Human lead exposure in England from approximately 5500 BP to the 16th century, AD. Sci Total Environ 318(1–3):45–58

Calabrese EJ, Baldwin LA (2001) The frequency of U-shaped dose-responses in the toxicological literature. Toxicol Sci 62:330–338

Carlon C, Swartjes F (2007) Analysis of variability and reasons of differences. In: Carlon C (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities, JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

Cornelis Ch, Berghmans P, Van Sprundel M, Van der Auwera J-C (2006) Use of the IEUBK model for determination of exposure routes in view of site remediation. Human Ecol Risk Assess 12(5):963–982

Cetin E, Odabasi M, Seyfioglu R (2003) Ambient volatile organic compound (VOC) concentrations around a petrochemical complex and a petroleum refinery. Sci Total Environ 312(1–3):103–112

Chen C, Huang GH, Chakma A (2003) Hybrid fuzzy-stochastic modeling approach for assessing environmental risks at contaminated groundwater systems. J Envir Eng 129(1):79–88

Chiodo LM, Covington Ch, Sokol RJ, Hannigan JH, Jannise J, Ager J, Greenwald M, Delaney-Black V (2007) Blood lead levels and specific attention effects in young children. Neurotoxicol Teratol 29(5):538–546

Chunilall V, Kindness A, Jonnalagadda SB (2006) Impact of coal mine dump contaminated soils on elemental uptake by *Spinacia oleracea* (spinach). J Environ Sci Health B 41(3):297–307

Davis S, Mirick DK (2005) Soil ingestion in children and adults in the same family. J Exposure Sci Environ Epidemiol 16:63–75

DEFRA, EA (2002) The contaminated land exposure assessment (CLEA) model: technical basis and algorithms. DEFRA (Department for Environment, Food and Rural Affairs) and EA (The Environment Agency), EA R&D publication CLR 10, Bristol, UK

Dourson ML, Anderson M, Erdreich LS, MacGregor J (2001) Using human data to protect the public's health. Regul Toxicol Pharmacol 33(22):234–256

Dybing E, Sanner T, Roelfzema H, Kroese D, Tennant RW (1997) A simplified carcinogenic potency index: description of the system and study of correlations between carcinogenic potency and species/site specificity and mutagenicity. Pharmacol Toxicol 80: 272–279

ECETOC (1995) Assessment factors in human health risk assessment. European centre for ecotoxicology and toxicology of chemicals, technical report no. 68, Brussels, Belgium

ECETOC (2001) Exposure factors sourcebook for European populations (with focus on UK data), technical report no. 79, Brussels, June, 2001, ISSN-0773-8072-79

ECHA (2008) Guidance on information requirements and chemical safety assessment. Chapter R.8 in Characterisation of dose (concentration)– response for human health, European chemicals agency. Available at http://reach.jrc.it/docs/guidance_document/information_requirements_r8_en.pdf?vers=20_08_08. Retrieved 30 Dec 2009

ECVAM (2009) http://ecvam.jrc.it/. Retrieved 24 Nov 2009

Edwards RD, Jurvelina J, Saarelad K, Jantunena M (2001) VOC concentrations measured in personal samples and residential indoor, outdoor and workplace microenvironments in EXPOLIS-Helsinki, Finland. Atmos Environ 35:4531–4543

EFSA (2005) Opinion of the scientific committee on a request from EFSA related to a harmonised approach for risk assessment of substances which are both genotoxic and carcinogenic. EFSA J 282:1–30

Environment Agency and DEFRA (2002) The contaminated land exposure assessment model (CLEA): technical basis and algorithms 2. Document prepared by the national groundwater and contaminated land centre of the environment agency R&D publication CLR10

Environment Agency (2008) Human health toxicological assessment of contaminants in soil. Science report SC050021/SR2, Environment agency, Bristol, UK. Available at http://www.environment-agency.gov.uk/static/documents/Research/scho0508bnqyee_2024094.pdf. Cited 15 Dec 2008

Falcó G, Bocio A, Llobet JM, Domingo JL, Casas C, Teixidó A (2004) Dietary intake of hexachlorobenzene in Catalonia, Spain. Sci Total Environ 322(1–3):63–70

Filipsson AF, Sand S, Nilsson J, Victorin K (2003) The benchmark dose method– review of available models, and recommendations for application in health risk assessment. Crit Rev Toxicol 33(5):505–542

Freeman NCG, Hore P, Black K, Jimenez M, Sheldon I, Tulve N, Lioy PJ (2005) Contributions of children's activities to pesticide hand loadings following residential pesticide application. J Exposure Anal Environ Epidemiol 15:81–88

Gilbert SG, Weiss B (2006) A rationale for lowering the blood lead action level from 10 to 2 μg/dL. NeuroToxicology 27(5):693–701

Government/Research Councils Initiative on Risk Assessment and Toxicology (1999) Physiologically-based pharmacokinetic modelling: a potential tool in risk assessment. Report of a workshop organised by the risk assessment and steering committee, Published by the institute for environment and health

Grosse SD, Matte TD, Schwartz J, Jackson R (2002) Economic gains resulting from the reduction in children's exposure to lead in the U.S. Environ Health Persp 110:563–569

Hawkins NC, Jayjock MA, Lynch J (1992) A rationale and framework for establishing the quality of human exposure assessments. Am Ind Hyg Assoc J 53(1):34–41

ILSI (1999) Significance of excursions of intake above the acceptable daily intake (ADI). Report of a workshop held in April 1998, organized by ILSI Europe acceptable daily intake task force and food chemical intake task force, ILSI Europe report series, International institute for life sciences Europe, Brussels, Belgium

Intawongse M, Dean JR (2006) Uptake of heavy metals by vegetable plants grown on contaminated soil and their bioavailability in the human gastrointestinal tract. Food Addit Contamin 23(1):36–48

International Programme on Chemical Safety (1999) Environmental Health Criteria no. 210– principles for the assessment of risks to human health from the exposure to chemicals. International programme on chemical safety (International programme on chemical safety), WHO, Genève, Switzerland

International Programme on Chemical Safety (2004) International programme on chemical safety risk assessment terminology – harmonization project document no. 1. International programme on chemical safety, WHO, Geneva, Switzerland

International Programme on Chemical Safety (2005) Harmonization project document no. 2– chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration – response assessment. International programme on chemical safety, WHO, Geneva, Switzerland

International Programme on Chemical Safety (2006) Environmental health criteria no. 237– principles for evaluating health risk in children associated with exposure to chemicals. International programme on chemical safety, WHO, Geneva, Switzerland

International Programme on Chemical Safety (2010) Environmental Health Criteria – Principles for modelling dose-response for the Risk Assessment of chemicals (Draft). International Programme on Chemical Safety, WHO, Geneva, Switzerland. http://www.who.int/ipcs/methods/harmonization/dose_response/en/index.html. Retrieved 18 August 2010

Iversen BS, Sabbioni E, Fortaner S, Pietra R, Nicolotti A (2003) Trace element reference values in tissues from inhabitants of the EU. XII, development of BioReVa program for statistical treatment. Sci Total Environ 302(1–3):1–12

JECFA (1987) Principles for the safety assessment of food additives and contaminants in food. Environmental health criteria 70. IPCS international programme on chemical safety in cooperation with the joint FAO/WHO expert committee on food additives (JECFA), Published under the joint sponsorship of the United Nations environment programme, the international labour organisation, and the world health organization in collaboration with the food and agriculture organization of the United Nations. http://www.who.int/pcs/jecfa/ehc70.html. Retrieved 30 Dec 2009

Johnson DL, Bretsch JK (2002) Soil lead and children's blood lead levels in Syracuse, New York, US. Environ Geochem Health 24:375–385

Kim B-H, Ikonomou MG, Lee S-J, Kim H-S, Chang Y-S (2005) Concentrations of polybrominated diphenyl ethers, polychlorinated dibenzo-p-dioxins and dibenzofurans, and polychlorinated biphenyls in human blood samples from Korea. Sci Total Environ 336(1–3):45–56

Koren HS, Seal E (2001) The ethical conduct of studies involving controlled human exposure to environmental pollutants. Human Ecol Risk Assess: Int J 7(6):1603–1606

Kouniali A, Cicolella A, Gonzalez-Flesca N, Dujardin R, Francois Gehanno J, Bois FY (2003) Environmental benzene exposure assessment for parent-child pairs in Rouen, France. Sci Total Environ 308(1–3):73–82

Krauss M, Wilcke W, Kobza J, Zech W (2002) Predicting heavy metal transfer from soil to plant: potential use of Freundlich-type functions. J Plant Nutr Soil Sci 165(1):3–8

Kulhánek A, Trapp S, Sismilich M, Janků J, Zimová M (2005) Crop-specific human exposure assessment for polycyclic aromatic hydrocarbons in Czech soils. Sci Total Environ 339(1–3):71–80

Lanphear BP, Matte ThD, Rogers J, Clickner RP, Dietz B, Bornschein RL, Succop P, Mahaffey KR, Dixon S, Galke W, Rabinowitz M, Farfel M, Rohde C, Schwartz J, Ashley P, Jacobs DE (2003) The contribution of lead-contaminated house dust and residential soil to children's blood lead levels: a pooled analysis of 12 epidemiologic studies. Environ Res 79(1):51–68

Leggett RW, Williams LR, Melo DR, Lipsztein JL (2003) A physiologically based biokinetic model for cesium in the human body. Sci Total Environ 317(1–3):235–255

Lehman AJ, Fitzhugh OG (1954) 100-fold margin of safety. Assoc Food Drug Official US Q Bull 18:33–35

Lemanek KL, Brown RT, Armstrong FD, Hood C, Pegelow Ch, Woods G (2002) Dysfunctional eating patterns and symptoms of pica in children and adolescents with sickle cell disease. Clin Pediatrics 41(7):493–500

Leroyer A, Hemon D, Nisse C, Auque G, Mazzuca M, Haguenoer J-M (2001) Determinants of cadmium burden levels in a population of children living in the vicinity of nonferrous smelters. Environ Res 87(3):147–159

Lindley F (2001) Creative ignorance. Human Ecol Risk Assess: Int J 7(6):1593–1601

Mackay D (2001) Multimedia environmental models– the fugacity approach, 2nd edn. CRC, Taylor and Francis, Boca Raton, FL, USA

Manta DS, Angelone M, Bellanca A, Neri R, Sprovieri M (2002) Heavy metals in urban soils: a case study from the city of Palermo (Sicily), Italy. Sci Total Environ 30(1–3):229–243

Maynard E, Thomas R, Simon D, Phipps C, Ward C, Calder I (2003) An evaluation of recent blood lead levels in Port Pirie, South Australia. Sci Total Environ 303(1–2):25–33

McConnell EE (2001) The value of human testing of pesticides. Human Ecol Risk Assess 7(6):1575–1581

McKone TE (1993) CalTOX: a multimedia total-exposure model for hazardous-wastes sites, Lawrence Livermore National Lab, CA (United States) technical report, UCRL-CR-111456-Pt.3

McMichael AJ, Woodward A (1999) Quantitative estimation and prediction of human cancer risk: its history and role in cancer prevention. In: Moolgavkar S, Krewski D, Zeise L, Cardis E, Møller H (eds) Quantitative estimation and prediction of human cancer risk, IARC scientific publication no. 131, International agency for research on cancer, Lyon, pp 1–10

Miekisch W, Schubert JK, Noeldge-Schomburg GFE (2004) Diagnostic potential of breath analysis—focus on volatile organic compounds. Clinica Chimica Acta 347(1–2):25–39

Mushak P (2003) Lead remediation and changes in human lead exposure: some physiological and biokinetic dimensions. Sci Total Environ 303(1–2):35–50

Otte PF, Lijzen JPA, Otte JG, Swartjes FA, Versluijs CW (2001) Evaluation and revision of the CSOIL parameter set. Proposed parameter set for human exposure modelling and deriving Intervention Values for the first series of compounds. RIVM report 711701021, RIVM, Bilthoven, the Netherlands

Pereira R, Ribeiro R, Gonçalves F (2004) Scalp hair analysis as a tool in assessing human exposure to heavy metals (S. Domingos mine, Portugal). Sci Total Environ 327(1–3):81–92

PETA (2009) http://www.peta.org/. Retrieved 24 Nov 2009

Provoost J, Cornelis C, Van Gestel G, Engels R (2004a) Basic information for risk assessments – part 3-H – Vlier-Humaan equations (in Dutch). OVAM, Mechelen, Belgium

Provoost J, Cornelis C, Van Gestel G, Engels R (2004b). Guidance for risk assessments – part 1-H – methodology for the derivation of soil remediation values (basic information on risk assessment) – part 1-H (in Dutch), OVAM, Mechelen, Belgium. Available from http://www.ovam.be. Retrieved 30 Dec 2009

Roberts SM (2001) Ethical issues in the use of data from testing of human subjects to support risk assessment. Human Ecol Risk Assess 7(6):1569–1573

Rodushkin I, Axelsson MD (2003) Application of double focusing sector field ICP-MS for multi elemental characterization of human hair and nails, Part III. Direct analysis by laser ablation. Sci Total Environ 305:23–39

Rosenkranz HS, Cunningham AR (2003) Environmental odors and health hazards. Sci Total Environ 313:15–24

Rossi E (2008) Low level environmental lead exposure – a continuing challenge. Clin Biochem Rev/Austr Assoc Clin Biochem 29(2):63–70

Ruby MV, Schoof R, Brattin W, Goldade M, Post G, Harnois M, Mosby DE, Casteel SW, Berti W, Carpenter M, Edwards D, Cragin D, Chappell W (1999) Advances in evaluating the oral bioavailability of inorganics in soil for use in human health risk assessment. Environ Sci Technol 33(21):3697–3705

Schroijen C, Baeyens W, Schoeters G, Den Hond E, Koppen G, Bruckers L, Nelen V, Van De Mieroop E, Bilau M, Covaci A, Keune H, Loots I, Kleinjans J, Dhooge W, Van Larebeke N (2008) Internal exposure to pollutants measured in blood and urine of Flemish adolescents in function of area of residence. Chemosphere 71(7):1317–1325

Sedki A, Lekouch N, Gamon S, Pineau A (2003) Toxic and essential trace metals in muscle, liver and kidney of bovines from a polluted area of Morocco. Sci Total Environ 317(1–3):201–205

Signorini S, Gerthoux PM, Dassi C, Cazzaniga M, Brambilla P, Vincoli N, Mocarelli P (2001) Environmental exposure to dioxin: the Seveso experience. Andrologia 32(4–5):263–270

Slob W (1999) Deriving safe exposure levels for chemicals from animal studies using statistical methods: recent developments. In: Barnett V, Stein A, Turkman KF (eds) Statistics for the environment 4: pollution assessment and control. Wiley, Chichester, pp 153–175

Slob W, Pieters MN (1998) A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. Risk Anal 18(6): 787–798

Standaert A, Van Holderbeke M, Cornelis C (2008) Modeling human exposure to cadmium and arsenic in the Northern Campine region. Epidemiology 19(6):S203

Speijers GJA (1999) Precision of estimates of an ADI (or TDI or PTWI). Regul Toxicol Pharmacol 30(2):S87–S93

Swartjes FA (1999) Risk-based assessment of soil and groundwater quality in the netherlands: standards and remediation urgency. Risk Anal 19(6):1235–1249

Swartjes FA (2002) Variation in calculated human exposure: comparison of calculations with seven European human exposure models. RIVM report 711701030, RIVM, Bilthoven, the Netherlands

Swartjes FA (2007) Insight into the variation in calculated human exposure to soil contaminants using seven different European models. Integr Environ Assess Manage 3(3): 322–332

Swartjes FA (2009) Evaluation of the variation in calculated human exposure to soil contaminants using seven different European models. Human Ecol Risk Asses 15:138–158

Swartjes FA, Baars AJ, Fleuren RHLJ, Otte PF (2004) *Risc limits for MTBE (Methyl tertiair-Butyl Ether) in soil, sediment, groundwater, surface water, drinking water and for drinking water preparation* (in Dutch), RIVM report 711701039, RIVM, Bilthoven, the Netherlands

Swartjes FA, d'Allesandro M, Cornelis Ch, Wcislo E, Müller D, Hazebrouck B, Jones C, Nathanail CP (2009) Towards consistency in Risk Assessment tools for contaminated site management in the EU. The HERACLES strategy from 2009 onwards. RIVM report 711701091, RIVM, Bilthoven, the Netherlands

Swartjes FA, Dirven-Van Breemen EM, Otte PF, Van Beelen P, Rikken MGJ, Tuinstra J, Spijker J, Lijzen JPA (2007) Towards a protocol for the site-specific human health risk assessment for consumption of vegetables from contaminated sites. RIVM report 711701040, RIVM, Bilthoven, the Netherlands

Tao S, Cui YH, Xu FL, Li BG, Cao J, Liu WX, Schmitt G, Wang XJ, Shen WR, Qing BP, Sun R (2004) Polycyclic aromatic hydrocarbons (PAHs) in agricultural soil and vegetables from Tianjin. Sci Total Environ 320(1):11–24

Teaf ChM (2008) Polycyclic aromatic hydrocarbons (PAHs) in urban soil: a florida risk assessment perspective. Int J Soil Sediment Water 1(2), Article 2

Thompson T (2002) Childhood blood lead screening, Anniston, Alabama, Agency for Toxic Substances and Disease Registry (ATSDR), January 2002, http://www.atsdr.cdc.gov/news/annistonreport.html. Retrieved 8 July 2010

Trapp S, Matthies M (1995) Generic one-compartment model for uptake of organic chemicals by foliar vegetation. Environ Sci Technol 29(9):2333–2338

Tsukahara T, Ezakia T, Moriguchia J, Furukib K, Shimboc S, Matsuda-Inoguchic N, Ikeda M (2003) Rice as the most influential source of cadmium intake among general Japanese population. Sci Total Environ 305:41–51

Turczynowicz L, Robinson N (2007) Exposure assessment modeling for volatiles—towards an australian indoor vapor intrusion model. J Toxicol Environ Health, Part A, 70(19):1619–1634

US Environmental Protection Agency (1992) Guidelines for exposure assessment (FRL-4129-5). US environmental protection agency, Washington DC, USA

US Environmental Protection Agency (1997) Exposure factors handbook. Volumes I–III. Office of research and development, US environmental protection agency, Washington, DC, EPA/600/P-05/002Fa

US Environmental Protection Agency (2000) Consideration of issues relating to the EPA's use of data derived from testing of human subjects. USEPA-SAB-EC-00-017, US environmental protection agency, Washington, DC, USA

US Environmental Protection Agency (2004) Risk assessment guidance for superfund volume I: human health evaluation manual (part e: supplemental guidance for dermal risk assessment)–final. US environmental protection agency, Washington, DC, EPA/540/R/99/005

US Environmental Protection Agency (2006) A framework for assessing health risks of environmental exposures to children. US environmental protection agency, Washington, DC, EPA/600/R-05/093F

US Environmental Protection Agency (2008) Child-specific exposure factors handbook, EPA/600/R–06/096F. National center for environmental assessment, Office of research and development, US environmental protection agency, Washington, DC 20460, September 2008

Van den Berg R (1991/1994/1995) Human exposure to soil contamination. A qualitative and quantitative analyses resulting in proposals for human health C screening values. Modified version of the original report Van den Berg 1991/Van den Berg 1994. RIVM report 725201011, RIVM, Bilthoven, the Netherlands

Van Holderbeke M, Cornelis Ch, Bierkens J, Torfs R (2008) Review of the soil ingestion pathway in human exposure assessment. VITO report 2008/IMS/R/0151

Vermeire T, Jager T, Janssen G, Bos P, Pieters M (2001) A probabilistic human health risk assessment for environmental exposure to dibutylphthalate. Human Ecol Risk Assess 7(6):1663–1679

Vermeire T, Stevenson H, Pieters MN, Rennen M, Slob W, Hakkert BC (1999) Assessment factors for human health risk assessment: a discussion paper. Crit Rev Toxicol 29(5):439–490

Vuori V, Zaleski RT, Jantunen MJ (2006) ExpoFacts—an overview of European exposure factors data. Risk Anal 26(3):831–843

Walker R (1999) The significance of excursions above the ADI: duration in relation to pivotal studies. Regul Toxicol Pharmacol 30(2):87–93

Walker JB, Seddon L, McMullen E, Houseman J, Tofflemire K, Corriveau A, Weber J-P, Mills C, Smith S, Van Oostdam J (2003) Organochlorine levels in maternal and umbilical cord blood plasma in Arctic Canada. Sci Total Environ 302(1–3):27–52

WHO (2005) Principles of characterizing and applying human exposure models. Harmonization project document no. 3

Wilkinson CF, Christoph GR, Julien E (2000) Assessing the risks of exposures to multiple chemicals with a common mechanism of toxicity: how to cumulate. Regul Toxicol Pharmacol 31:30–43

Wilson NK, Chuang JC, Lyu Ch, Menton R, Morgan MK (2003) Aggregate exposures of nine preschool children to persistent organic pollutants at day care and at home. J Exposure Anal Environ Epidemiol 13:187–202

WWF (2003) WWF-UK National biomonitoring survey 2003. World Wildlife Fund. Available via http://www.wwf.org.uk/filelibrary/pdf/biomonitoringresults.pdf. Retrieved 2 Dec 2008

Yoshida K, Ikeda S, Nakanishi J (2000) Assessment of human health risk of dioxins in Japan. Chemosphere 40:177–185

Chapter 6
Exposure Through Soil and Dust Ingestion

Johan Bierkens, Mirja Van Holderbeke, Christa Cornelis, and Rudi Torfs

Abstract The current chapter discusses soil and dust ingestion, a potentially important pathway for non-dietary oral exposure, especially for children. Starting from clear definitions on what is meant by soil and dust and how they interrelate, it explores the several approaches that have been used to derive estimates of soil and dust ingestion rates. It concludes that tracer methodology studies, with all their limitations and uncertainties, probably provide the most adequate estimates of soil ingestion rates. However, these studies are limited to short-term estimates and do not distinguish between soil and dust. Hand-loading studies can be designed so that information is collected for the micro-environments we are interested in, but the interpretation requires assumptions about transfer parameters, which may introduce substantial uncertainty. Biokinetic modelling studies only provide rough estimates or ranges of estimates, but are nevertheless useful as a complementary line of evidence. From a comparison of these approaches it is concluded that average soil and dust ingestion rates for children are below 100 mg/d and most likely around 50 mg/d, a conclusion that is confirmed by the most recent evaluations published in literature. Still, good estimates of site-specific soil and dust ingestion rates for Europe are lacking due to paucity of data on age-related time activity patterns, transfer factors and intrinsic differences in children's behaviour.

Contents

J. Bierkens (✉)
Flemish Institute of Technological Research (VITO), Mol, Belgium
e-mail: Johan.bierkens@vito.be

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_6,

6.1 Introduction

6.1.1 General Aspects

Ingestion of soil and dust is a potentially important pathway of exposure to environmental contaminants, especially for immobile and semi-volatile contaminants such as metals, PAHs and POPs. Ingestion of soil can be a special concern for young children who may, per kilogram body weight, ingest up to 10 times more soil than adults (Egeghy et al. 2007). However, in assessing the relative contribution of ingestion of soil and dust to overall oral exposure, we need to be aware that, depending on the agent, the oral bioavailability of contaminants in soil can be lower compared to e.g. the oral bioavailability in food.

Due to their different behaviour, young children show exposure patterns distinct from those of adults. Young children play close to the ground and may ingest significant quantities of outdoor soil, and soil and dust from indoor surfaces and carpets. Soil and dust ingestion may result from various behaviours including, but not limited to, mouthing (putting objects or hands in the mouth), eating dropped food, or consuming soil or dust directly. A special case of soil ingestion is the so-called Soil-pica, which is the recurrent ingestion of unusually high amounts of soil, i.e., on the order of 1,000–5,000 mg/d. *Geophagy* is the intentional ingestion of soil material and is usually associated with cultural practices (ATSDR 2001).

As soil and dust ingestion constitutes one of the major exposure routes for children in case of contaminated sites, and to a lesser extent for adults, it is of utmost importance to derive accurate soil and dust ingestion values for Risk Assessment.

6.1.2 Defining Soil and Dust

Soil consists of particles of unconsolidated mineral and/or organic matter from the earth's surface that are located outdoors, or are used indoors to support plant growth. It includes particles that have settled onto outdoor objects and surfaces (US-EPA 2008). House dust is a heterogeneous mixture of tracked-in or re-suspended soil and non-soil particles, including clothing fluffs, atmospheric deposited particulates, hair, fibres (artificial and natural), moulds, pollen, allergens, bacteria, viruses, arthropods, ash, soot, animal fur and dander, smoke, skin particles, cooking and heating residues, and building components.

In this chapter we therefore use the following definitions:

- Soil: outdoor soil and soil used indoors to support plant growth, including particles that have settled onto outdoor objects and surfaces;
- Dust or house dust: settled indoor material, consisting of soil-derived and non-soil-derived materials.

On average, house dust includes 30–70% soil material, indicating that contaminated soil can lead to contaminated house dust (Oomen and Lijzen 2004).

The fraction of exterior soil in interior dust has been examined by various researchers. Oomen and Lijzen (2004) and Cornelis and Swartjes (2007, see Table 6.1) give a good overview of the fraction of exterior soil in indoor dust derived from different relevant studies.

From Table 6.1 it can be concluded that estimates of the contribution of soil to house dust range from 8 to >80%, depending on a wide variety of site-specific factors and methodological approaches. Cornelis and Swartjes (2007) recommend the use of 50% exterior soil in interior dust for a residential land use. For the land use "living without a garden" they propose a fraction of 25% exterior soil in interior dust. The default value recommended by US EPA (1998) is 70% exterior soil in interior dust.

The distinction between soil and house dust (including tracked-in soil) may be very important in Risk Assessment. House dust is generally composed of finer particles than soil. Particles with a diameter exceeding 150 μm represent 80% of

Table 6.1 Overview of the fraction of exterior soil in indoor dust derived from different studies

Study	Percent of exterior soil in house dust
Hawley (1985) in Oomen and Lijzen (2004)	>80
Thornton (1985) in Oomen and Lijzen (2004)	20
Camann and Harding (1989) in Oomen and Lijzen (2004)	50
Fergusson and Kim (1991) in Oomen and Lijzen (2004)	32–50
Calabrese and Stanek (1992) in Oomen and Lijzen (2004)	31
Sterling (1998) in Oomen and Lijzen (2004)	37 (based on particle volume weighting) 8 (based on particle concentration weighting)
Rutz, Valentine, Eckart and Yu (1997)	20, 30
Trownbridge and Burmaster (1997)	P50: 41 Arithmetic mean: 44.5 SD: 16.8 %
Hager (2005)	P50: 26–27 P25: 16–14 P75: 44–42 Arithmetic mean: 42–37

Cornelis and Swartjes (2007)

external soil, but only 50% of house dust. In addition to being more mobile, fine particles adhere more effectively to the skin, thus increasing the potential for exposure. Furthermore, the degrading or ageing processes that are effective at removal of contaminants outside do not operate or they operate at reduced efficiency indoors compared to outdoors, because they are protected from sunlight, rain, temperature extremes, and microbial action (Egeghy et al. 2007). Contaminant concentrations in house dust are often higher than in exterior soil. This could be due to enrichment in the finer fractions and lack of removal processes. Another explanation may be the possible presence of indoor sources of contaminants, which can contribute significantly to overall concentrations in house dust. Given the importance of tracked-in soil on indoor contaminant concentrations, exposure to soil-derived dust should be included in Risk Assessment in regard to contaminated sites.

6.1.3 Calculating Exposure Through Ingestion of Soil and Dust

Exposure to contaminants via ingestion of house dust and soil can be calculated according to the following equations,

$$D_{\text{soil ingestion}} = \frac{IR_{\text{soil/dust}} * F_{\text{soil}} * RBA_{\text{soil}} * C_{\text{soil}}}{BW} \tag{6.1}$$

$$D_{\text{dust ingestion}} = \frac{IR_{\text{soil/dust}} * (1 - F_{\text{soil}}) * RBA_{\text{dust}} * C_{\text{dust}}}{BW} \tag{6.2}$$

where

$D_{\text{soil ingestion}}$ – exposure through soil ingestion ($\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$)
$D_{\text{dust ingestion}}$ – exposure through dust ingestion ($\text{mg}\cdot\text{kg}_{\text{bw}}{}^{-1}\cdot\text{d}^{-1}$)
$IR_{\text{soil/dust}}$ – combined ingestion rate of soil and dust (kg/d^{-1})
RBA_{soil} – relative oral bioavailability in the human body for contaminants in soil (–)
RBA_{dust} – relative oral bioavailability in the human body for contaminants in dust (–)
$F_{\text{oral\_soil}}$ – fraction of soil and dust ingestion which is soil (–)
C_{soil}– contaminant concentration in soil ($\text{mg}\cdot\text{kg}_{\text{dw}}{}^{-1}$)
C_{dust} – contaminant concentration in dust ($\text{mg}\cdot\text{kg}_{\text{dw}}{}^{-1}$)
BW – body weight (kg)

The derivation of soil and dust ingestion rates is discussed in detail in this chapter. If, depending on the studies used, a separate estimate is available for the soil ingestion rate and for the dust ingestion rate, these values can replace the product of $IR^{*}F$.

Relative oral bioavailability in the human body of contaminants from soil allow transformation of the exposure from soil and dust to an exposure that can be compared to the Toxicological reference value or that can be summed up as exposure from other oral exposure pathways.

Concentrations of contaminants in dust can be measured or can be estimated from concentrations in soil and the fraction of soil in indoor dust. When making measurements, the risk assessor should be aware that the concentrations observed are not necessarily solely from a soil origin. Distinguishing between soil-derived and non-soil-derived concentrations is often a challenge in site-specific Risk Assessments. Estimating the concentration in dust that is derived from soil is uncertain as well, as enrichment and dilution processes can be highly variable.

6.2 Quantification of Soil and Dust Ingestion Rates

Estimates of soil and dust ingestion can be based both on *Macro and Micro activity* data. The Macro activity approach is based on tracer studies. The Micro activity approach focuses on hand-loading studies.

In the Macro activity approach exposure is estimated individually for each of the micro-environments where a child spends time, e.g., a living room, courtyard, et cetera, and for each Macro activity a child conducts within that micro-environment, e.g., playing, reading, et cetera). In this case exposure is considered as the aggregate exposure of a series of contacts with the contaminated medium (soil or dust). In the Micro activity approach exposure is modelled as a series of discrete transfers resulting from each contact with the contaminated medium.

Other, far less used methodologies can be applied to estimate the amount of soil and dust ingested by children, such as the Biokinetic model comparison methodology, the Lead isotope ratio methodology, the Survey response methodology, and the Empirical relations methodology (US-EPA 2008; Van Holderbeke et al. 2008).

6.2.1 Tracer Element Methodology

The *tracer element methodology* attempts to quantify soil ingestion rates outdoors and soil tracked into the indoor environment. In this methodology, both samples of soil from children's residences and samples of the children's excreta (faeces and sometimes also urine) are analysed for the presence and quantity of soil borne tracer elements, mainly, aluminium (Al), silicon (Si), titanium (Ti), and yttrium (Y). Ideally, tracer elements are not metabolised or absorbed from the gastrointestinal tract in significant quantities, and their presence in soil, faeces and urine can be used to estimate soil ingestion rates using a mass balance methodology. This methodology assumes that the amount of soil (outdoors and tracked indoors) ingested in grams equals the quantity of a given tracer element present in the faeces and urine (in mg), minus the quantity of that tracer element, present in food and medicine (mg), divided by the tracer element's soil concentration (in mg/g) (US-EPA 2008). Some authors substitute tracer element soil concentrations with tracer element dust concentrations, which results in a dust ingestion estimate. The tracers in soil are present in soil outdoors and in soil in house dust and therefore the amount of non-food quantity ingested is a combination of soil and dust. The true soil and/or dust ingestion is probably somewhere between the estimates based on soil concentrations and the estimates based on dust concentrations (US-EPA 2008). By making assumptions on the contribution of soil and dust to the overall ingestion rate, some authors attempted to estimate a combined soil and dust ingestion rate.

Initially the tracer element methodology did not account for the contribution of tracer elements from non-soil sources (food, medications, and non-food sources such as toothpaste), e.g., Binder et al. (1986). Later studies generally account for tracer element contributions from these non-soil sources (Calabrese et al. 1997; Davis et al. 1990).

A number of key studies have been conducted to estimate soil ingestion rates using the tracer element methodology (Binder et al. 1986; Calabrese et al. 1989, 1997; Clausing et al. 1987; Davis et al. 1990; Van Wijnen et al. 1990). In the second instance, the experimental data of many of these studies were given new interpretations (Calabrese and Stanek 1992, 1995; Calabrese et al. 1996; Davis and Mirick 2006; Sedman and Mahmood 1994; Stanek and Calabrese 1995a, b, 2000; Stanek et al. 1999, 2001a; Thompson and Burmaster 1991). Most of these studies have estimated soil ingestion rates for children. Data on ingestion rates for adults are scarce. Chapter 5 of US EPA's Child-Specific Exposure Factors Handbook (2008) and Van Holderbeke et al. (2008) present a good overview of the relevant studies for children. Details on the available studies are presented in Tables 6.2 and 6.3.

Differences exist between the above listed studies in respect to the correction that is made for ingestion of tracers from non-soil related sources (food, toothpaste, et cetera) and the tracers that are measured. In addition, the interpretation of the results between studies differs. Some authors use the *Limited Tracer Method* (LTM), in which for every sample the lowest of all tracers is used (e.g., Binder et al. 1986; Clausing et al. 1987; Van Wijnen et al. 1990). In contrast, Calabrese et al. (1997) estimated the soil ingestion rate by each tracer element using the *Best Tracer Method* (BTM), which allows for the selection of the most recoverable tracer on the basis of the food/soil ratio for a particular group of subjects (Stanek and Calabrese 1995b). Some authors adjust their estimated soil ingestion rates to account for the potential contribution of tracer elements found in house dust, as well as in soil (Davis et al. 1990).

In comparison with children, data on soil ingestion for adults are much more limited. Details on the relevant studies are given in Table 6.3.

Soil and dust ingestion rates derived from tracer studies are subject to uncertainties, e.g., the experimental uncertainty in the measurements by only partial sampling of the food and faeces, or transit time differences, using standard stool weights and not actual stool weights. Another uncertainty is the uptake of tracers by the human body: the assumption that tracer elements are not absorbed is not entirely true. The methodology assumes that children ingest predominantly soil from their own yards. In reality, however, children ingest also soil from other locations. Other limitations are inherent to soil sampling (e.g., homogeneity of the soil samples) and laboratory analytical techniques. In addition, the true soil and/or dust ingestion rate is probably somewhere between the estimates based on soil concentrations and the estimates based on dust concentrations.

Table 6.2 Overview of the relevant tracer studies on soil ingestion rates by children

References	Data used	Number of children	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/ climate	Duration	Method
Davis et al. (1990)	Davis et al. (1990)	101	2–7	61	42	Three-city area in south-eastern Washington State	Primarily summer/desert climate: cold in winter, very hot in summer and very low precipitation year-round	4-day period	Correction for the contribution of tracer elements from food, vitamins and medications, evaluation of the extent to which differences in tracer concentrations in house dust and yard soil impact estimated soil ingestion rates
Davis and Mirick (2006)	Subset of Davis et al. 1990	12		37	30			11 consecutive days	Correction for the contribution of tracer elements from food, vitamins and medications, daily diary of activities for 4 consecutive days during the study period, collection of soil and house dust samples

Table 6.2 (continued)

References	Data used	Number of children	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/ climate	Duration	Method
Calabrese et al. (1997)	Calabrese et al. (1997)	64	1–3	66	20	EPA Superfund site: contaminated site in Anaconda, Montana	September/cool, dry summers and cold, snowy winters	7 consecutive days	Duplicate samples of meals, beverages, and medicines and vitamins; toothpaste, baby cornstarch, diaper rash cream, and soap were provided; Best Tracer Method (best four trace elements), collection of soil and house dust samples
Stanek et al. (2001b)				31	24				Idem Calabrese et al. (1997) + estimation of long-term (annual) average soil ingestion distribution
Calabrese et al. (1996)									Idem Calabrese et al. (1997) + soil tracer concentrations by particle size of <250 µm in diameter
Stanek et al. (1999)									Idem Calabrese et al. (1997) + soil tracer concentrations by particle size of 100–250 µm and to particle sizes of 53 to < 100 µm

Table 6.2 (continued)

References	Data used	Number of children	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/ climate	Duration	Method
Stanek and Calabrese (2000)				31	17				Idem Calabrese et al. (1997) + assumption of a lognormal distribution for the soil ingestion estimates to predict average soil ingestion over a longer time period
Binder et al. (1986)	Binder et al. (1986)	59	1–3	108		Residents living near a lead smelter in East Helena, Montana.	Summer/modified continental climate with warm, dry summers and moderately cold winters; April through July is the rainy season, while late summer, fall, and winter are quite dry.	3-day period	Analysis of excreta and soil samples from the children's yards, no distinguish between ingestion of soil and house dust; no correction for the presence of the tracer elements in ingested foods or medicines
Thompson and Burmaster (1991)				91					Idem Binder et al. (1986) + using the actual stool weights of the study participants instead of the assumed stool weights

Table 6.2 (continued)

References	Data used	Number of children	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/ climate	Duration	Method
Clausing et al. (1987)	Clausing et al. (1987)	18 (school) + 6 (hospitalised, bedridden) children	2–4	56	82	Children attending the same nursery school in the Netherlands	Temperate climate	5-day period	No correction for tracer elements in foods or medicines; hospitalised, bedridden children served as a control group, representing children who had very limited access to soil
Calabrese et al. (1989)	Calabrese et al. (1989)	64	1–4	131	26	Greater Amherst, MA	Cold winters, moderate summers	8 days spread over 2 weeks	Correction for tracer elements in food, beverages, medicines, and vitamins; collection of soil and dust samples
Stanek and Calabrese (1995a)				179	45				Idem Calabrese et al. (1989) + linking the physical passage of food and faecal samples; distribution of soil ingestion projected over a period of 365 days
Calabrese and Stanek (1992)									Idem Calabrese et al. (1989) + estimation of the amount of outdoor soil in indoor dust

Table 6.2 (continued)

References	Data used	Number of children	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/ climate	Duration	Method
Calabrese and Stanek (1995)				122					Idem Calabrese et al. (1989) + reduction of the magnitude of error in the individual trace element ingestion estimates
Stanek and Calabrese (1995b)	Calabrese et al. (1989); Davis et al. (1990)	229	2–4 in Calabrese et al. (1989) and 4.7 in Davis et al. (1990)	113	37	Data from two studies for children, one from Washington and one from Massachusetts			Best Tracer Method (BTM)
Sedman and Mahmood (1994)				195					Mass balance approach: dividing the excess tracer intake (i.e., quantity of tracer recovered in the faeces in excess of the measured intake) by the average concentration of tracer in soil samples from each child's dwelling

Table 6.2 (continued)

References	Data used	Number of children	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/ climate	Duration	Method
Van Wijnen et al. (1990)	Van Wijnen et al. (1990)	292/187 children attending day-care centres; 78 children at campgrounds; 15 hospitalised children	1–5	69/120	114/160	The Netherlands: day-care centres, campgrounds, hospital	Summer (June–July and August–September)/ temperate climate	3–5 days	Limiting tracer method (LTM); no correction for tracer elements in foods or medicines; soil intake was estimated for three different populations of children: one expected to have high intake, one expected to have "typical" intake, and one expected to have low or background-level intake

*To calculate mean and median values only those data that met the criteria specified by Van Holderbeke et al. (2008) were used, e.g. no values on Ti are included and all data were corrected for background whenever background values are provided in the original paper

Table 6.3 Overview of the relevant tracer studies on soil ingestion rates by adults

Reference	Data used	Number of adults	Age (years)	Mean ingestion rate (mg/d) (*)	Median ingestion rate (mg/d) (*)	Area	Season/climate	Duration	Method
Calabrese et al. (1990)	Calabrese et al. (1989)	6	25–41	39	30	Greater Amherst, MA	Cold winters, moderate summers	3 weeks	Administration of known amounts of soil from 0 to 500 mg/d; collection of duplicate meal samples
Stanek et al. (1997)		10		44–49	18			4 weeks	Administration of known amounts of soil from 0 to 500 mg/d
Davis and Mirick (2006)	Subset of Davis et al. (1990)	Mothers and fathers of 19 families		23–92	0.2–23	Three-city area in south-eastern Washington State	Primarily summer/desert climate: cold in winter, very hot in summer and very low precipitation year-round	11 days	Collection of food/liquid, excreta, and soil/dust samples, mass balance approach

* To calculate mean and median values, only those data that met the criteria specified by Van Holderbeke et al. (2008) were used, e.g., no values on Ti are included and all data were corrected for background whenever background values are provided in the original paper

An important limitation of tracer studies is that it is not possible to distinguish between soil and dust ingestion. However, this distinction is important because of differences in concentrations and bioavailability and for adequate design of Risk Management solutions. Estimates that were made by some authors (e.g., Davis et al. 1990) provide assumptions about how the ingestion rate is distributed over soil and dust, but data to support these assumptions are limited. The contribution of soil and dust in the overall ingestion rate may be not only related to the time children spent outdoors and indoors, but also to the playing pattern and the higher oral availability in the human body of soil material compared to indoor dust material.

6.2.2 Alternative Approaches for Estimating Soil and Dust Ingestion

Estimates of soil and dust ingestion based on data of *hand-loading and mouthing behaviour* of children require detailed knowledge on time activity patterns, frequency of hand to mouth contact, age specific hand-mouth contact area, soil/dust hand-loading and an estimate of transfer efficiencies, i.e. the degree to which soil or dust adheres to skin and the percentage of soil adhering to the hand that is ingested during a single hand to mouth contact. An advantage of hand-loading studies is that ingestion rates can be derived separately for outdoor and indoor activities and thus for soil and dust ingestion

Finley et al. (1994) and Van Holderbeke et al. (2008) give an overview of studies that have been conducted to estimate soil/dust hand-loading for children of different ages, locations (indoor and outdoor) and sampling methods (gravimetric and wipe method). Finley et al. (1994) interpreted these studies using different assumptions such as surface area of hands, percentage of children's hand covered with soil, efficiency of the methods and concentrations in soil and house dust. From these studies it can be concluded that dermal hand-loadings for children generally range from 0.1 to 1.0 mg soil or dust/cm^2.

It is difficult to quantify the hand-to-mouth transfer of soil, dust or contaminants. Not only the duration and type of contact (licking, sucking, et cetera), but also the polarity and affinity of contaminants to saliva and skin are important. Kissel et al. (1998) reported the results of a laboratory-based examination of hand-to-mouth transfer of soil. The mass fraction of total soil load on the hand recovered from mouths following thumb sucking, palm licking or mouthing three fingers above the first knuckle ranged from 10.1 to 21.9%. It should be noticed however that this study was conducted with adults and does not take into account the difference in hand/finger proportions that exists between adults and toddlers. These data provide the soil mass transferred to the mouth divided by the soil mass on the entire hand, rather than the part of the hand that is in contact with the mouth. Therefore, they can be used to estimate an extreme lower bound of the hand-to-mouth transfer efficiency factor (OEHHA 2007). Most authors investigated the transfer efficiency of contaminants such as pesticides (Camann et al. 2000) or riboflavin (Cohen Hubal et al. 2005). Values in these papers range from 5 to 100% extraction. Nearly all authors use 50% as the efficiency of the hand-to-mouth transfer.

Hand-to-mouth frequency data are limited and difficult to collect. Few published studies report hand-to-mouth frequency data. Different data collection approaches are used, such as videotaping versus real-time observation, data analysis and reporting methods. Studies differ also by ages of children, locations and definitions of mouthing. Xue et al. (2007) pooled and reanalysed indoor and outdoor mouthing frequency data. The results of this meta-analysis are consistent with the results from other studies reported in the literature. Most authors conclude that age and location (indoor or outdoor) are important for hand-to-mouth frequency: as age increases, frequencies decrease (Freeman et al. 2001; Tulve et al. 2002). Hand-to mouth frequency is also higher indoors than outdoors (AuYeung et al. 2004; Freeman et al. 2001) and some other authors (AuYeung et al. 2004; Freeman et al. 2001) observe higher frequencies for girls than for boys.

Most authors assume that each hand-to-mouth event equals a contact area of 3 fingers (Kuusisto et al. 2007; US-EPA 2001). Depending on the age, the contact area ranges between 11 and 33 cm^2 from infant to 12 year old children (Kuusisto et al. 2007).

The results from hand-to-mouth transfer studies typically provide hourly ingestion rates of soil and/or house dust which can be compared with the daily ingestion rates derived from tracer element studies when information exists on how children have spent their time indoors and outdoors. As the location and the type of activity children are involved in (indoor or outdoor activities; passive or dynamic play) determine the results from hand-loading studies, care should be taken when comparisons with tracer studies are made. Hand-loading studies can theoretically provide the most accurate soil and dust ingestion values for humans and allow distinction between soil (outdoor) and dust (indoor) ingestion. However, since many variables have to be taken into account, the resulting soil and dust ingestion rates show a rather high variation.

The *Biokinetic model comparison methodology* is based on the validation of exposure models for contaminants which have soil or dust ingestion as one of the major exposure routes and in which modelled and measured internal exposure can be compared (Van Holderbeke et al. 2008; US-EPA 2008). Only studies that meet the following criteria can be used to estimate soil/dust ingestion rates:

- the contaminant should contribute significantly and preferably dominantly to the exposure via the soil/dust ingestion pathway;
- the biomarker for model validation should relate to actual exposure;
- an adequate pharmacokinetic model, linking external or absorbed dose to the biomarker, should be available.

Case studies that meet these criteria are mostly available for lead and children's exposure. Some studies for arsenic are also available. The methodology is at present limited to rough estimates, due to the uncertainty with regard to oral bioavailability in the human body (if not measured), contribution of soil versus dust and the representativeness of the soil and dust concentrations measured. Simulations can be performed to obtain an optimal fit based on modifying the soil/dust ingestion rates and the contribution of soil and dust. However, in order to be able to do this the raw data records from the original studies are required, data that are often not available. Equally, age dependency of the contribution of soil and dust and of ingestion rates can play a role and is often not considered in these studies.

In the *survey response methodology*, responses to survey questions regarding soil and dust ingestion, e.g., soil and dust ingestion behaviour, frequency and quantity, are analysed (US-EPA 2008). In these studies, children or children's caretakers are surveyed. Different data collection methodologies can be used, e.g., on the basis of personal interviews or mailed questionnaires. Each collection methodology has specific limitations, which results in either positive or negative response bias. Respondents' perception of a correct answer, question wording difficulties, lack of understanding of definitions and terms used, language and dialect differences, etc can cause bias (US-EPA 2008).

According to the *lead isotope ratio methodology*, lead isotope ratios in children's urine or blood, water, food, and house dust are measured and compared to infer sources of lead exposure that may include dust or other environmental exposures (US-EPA 2008).

Another approach that might be used to derive soil and dust ingestion rate values is the use of *empirical relations between contaminant levels in the environment and Biomonitoring data*. In this approach, the slope between the biomarker and the concentration in soil or dust is used to calculate the ingestion rate. As this approach is based on statistical regression analysis and not on mechanistic processes, regressions that account for as many variables as possible (e.g., dust, soil, food, air, age, socio-economic factors) are preferred, so as to obtain a slope which is optimally related to soil or dust ingestion.

6.2.3 Soil and Dust Ingestion Rates for Children and Adults

Most authors derive soil ingestion rates from a selection of tracer studies. However, as mentioned above, also other methodologies can be applied to estimate the amount of soil and dust ingested by children.

Until recently, the US EPA recommended soil ingestion values for children from age 1–7 years on the basis of some key studies, i.e., studies with a superior experimental design or with a particular attribute needed for the recommendations (US-EPA 2006). The arithmetic mean ingestion rates ranged from 38 mg/d to 193 mg/d with a weighted average of 90 mg/d for soil ingestion and 106 mg/d when it was considered that a portion of the soil ingested comes from dust. These estimates are based on weighted averages using aluminium and silicon as tracers. Also for US EPA these tracer elements were considered the most reliable, based on a review of the current literature. Results obtained using titanium as a tracer were not considered. The reason for this is that in some children's diet titanium is present in large quantities and titanium dioxide pigment in paints may provide an additional source of titanium exposure through the ingestion of paint chips or dust (Davis et al.1990).

Rounding our figures, 100 mg/d was considered the best estimate of the arithmetic mean soil ingestion for children under 7 years of age. Over the period of study, the 95th percentile ingestion rates ranged from 217 to 449 mg/d with an arithmetic mean of 236 mg/d for soil ingestion and 449 mg/d when both soil and dust are considered. Rounding off, the recommended 95th percentile soil ingestion rate for children is 400 mg/d based on soil and dust ingestion. In its revision of the Child-specific Exposure Factors Handbook (US-EPA 2008), US EPA re-adjusted their recommendations for values for daily soil, dust and soil/dust ingestion to much lower ingestion values, see Table 6.4. The revision relies on tracer studies and one Biokinetic model comparison study. Ingestion values from tracer studies using soil concentrations are interpreted as ingestion of soil outdoors and from soil used indoors to support plant growth. Soil particles that have been tracked into the indoor environment from outdoors are not included. Estimated ingestion values using dust concentrations are used for the estimation of dust ingestion, including soil particles tracked to the indoor environment. To derive the soil and dust ingestion values in Table 6.4, the US EPA assumes that the relative proportions of soil and dust ingested are 45% soil and 55% dust, respectively.

US-EPA (2008) also provides an estimate for soil-pica (1,000 mg/d), which is based on their definition (ingestion of high amounts of soil in the order of 1,000–5,000 mg/d and the finding that in ingestion studies, the high-end of the intakes exceed that quantity).

US-EPA (1997) recommends a CTE (Central Tendency Exposure) soil ingestion value for adults of 50 mg/d, although they claim that this recommendation is highly uncertain. Considering the uncertainties in the central estimate, a recommendation for an upper percentile value would be inappropriate.

Zartarian et al. (2005) conducted an analysis of soil ingestion rates from selected tracer studies to be used as input for the SHEDS-WOOD (Stochastic

Table 6.4 Recommended values for daily soil, dust and soil/dust ingestion by children) (US-EPA 2008)

	Soil[a]			Dust[b]	Soil + dust
	Central tendency (mg/d)	Upper percentile		Central tendency (mg/d)	Central tendency (mg/d)
Age group		Soil-pica (mg/d)	Geophagy (mg/d)		
6 to <12 months	30	–	–	30	60
1 to <6 years	50	1,000	50,000	60	100[c]
6 to <21 years	50	1,000	50,000	60	100[c]

– No recommendation

[a] Includes soil and outdoor settled dust

[b] Includes indoor settled dust only

[c] Total soil and dust ingestion rate is 110 mg/d, rounded to one significant figure it is 100 mg/d

Human Exposure and Dose Simulation) model. Soil ingestion rate estimates were derived using data from Calabrese's Amherst and Anaconda studies (by Stanek and Calabrese (2000) and Stanek et al. (2001a)). Data statistics from both of these studies were used to fit distributions of soil/dust ingestion rates (Stanek et al. 2001b). The statistical distributions generated for variability and uncertainty distributions relied upon two tracers only, aluminium and silicon, in estimating the parameters of the lognormal variability and uncertainty distributions. Using Monte-Carlo sampling, values from the fitted distribution were sampled. The sampled values were separated into those values under 500 mg/d and values that exceeded 500 mg/d. The model assumes that soil ingestion values that exceed 500 mg/d are representative for pica behaviour. The soil ingestion rate distribution for non-pica behaviour children and children exhibiting pica behaviour are shown in Table 6.5.

The age-specific default soil/dust ingestion rates recommended for use in the Integrated Exposure Uptake Biokinetic (IEUBK) model for lead in children range from 85 to 135 mg/d and are illustrated in Table 6.6. These values are considered representative of arithmetic mean daily intake rates. The default soil and dust ingestion rates are based on the tracer studies of Binder et al. (1986), Clausing et al. (1987), Calabrese and Stanek (1992) Calabrese et al. (1989), Van Wijnen et al. (1990), and Davis et al. (1990).

Van Holderbeke et al. (2008) conducted an analysis of soil ingestion rates from tracer studies, hand-loading studies, Biokinetic model comparison and empirical relations between contaminant levels in the environment and Biomonitoring data. The authors selected tracer studies fulfilling the following requirements:

- data were corrected for intake of tracers from sources other than soil;
- only the most recently published re-analysed results were selected (published and re-analysed by the same authors);
- data on titanium were left out, because this was considered the least reliable tracer.

Table 6.5 The soil ingestion rate distribution for non-pica behaviour children and children exhibiting pica behaviour as input for the SHEDS-WOOD model

Scenario	Variability distribution	Arithmetic mean	St. dev.	Median	P25	P75	P95	P99	Resources
Typical child	Lognormal (31.4) < 500 mg/d	60.6	80.5	29.8	11.9	73.4	236	402	Stanek and Calabrese (2000); Stanek et al. (2001a)
Pica child	Lognormal (31.4) > 500 mg/d	962	758	735	590	1046	2130	3852	ATSDR (2001)

Zartarian et al. (2005)

Table 6.6 The age-specific default soil/dust ingestion rates recommended for use in the Integrated Exposure Uptake Biokinetic (IEUBK) model for lead in children

Age groups (years)	IEUBK-model defaults (mg/d)
0–1	85
1–2	135
2–3	135
3–4	135
4–5	100
5–6	90
6–7	85

US EPA (1999)

Ingestion values from tracer studies using soil concentrations were interpreted as ingestion of soil outdoors, soil used indoors to support plant growth and from soil particles that have been tracked into the indoor environment. This is a different interpretation than in US EPA (2008). Results based on Calabrese et al. (1989) were left out, because Calabrese claimed in 2003 that these data overstate the soil ingestion rate. The result of the re-analysis of literature data was a set of selected soil ingestion values that met the above-mentioned criteria. Mean values (± standard deviation) were calculated, and these data were subsequently used for bootstrapping. Parametric bootstrapping was performed on the lognormal distribution fitted through the arithmetic mean from these studies. The distribution results in a mean soil (combined outdoors and tracked indoors) ingestion rates of 63 mg/d (95% = 81 mg/d) and a 95th percentile of 97 mg/d (95% = 124 mg/d; Table 6.7). The authors applied a model developed by Bierkens and Cornelis (2006) that derives ingestion rates for soil indoor and outdoor and dust from tracer study data, using information on the fraction of soil in indoor dust and the ratio of ingestion indoors compared to ingestion outdoors. This model results in a daily soil and dust ingestion rate of 87 mg/d for children, which is slightly lower than the 100 mg/d proposed in US EPA 2008.

Van Holderbeke et al. (2008) estimated the mean soil ingestion for adults based on the available tracer studies. The arithmetic mean daily soil (outdoors/tracked-in) ingestion rates for adults range between 5 and 92 mg/d, taking into account the background exposure. The arithmetic mean of all these values is 46 mg/d. When

Table 6.7 Statistical descriptors of the distributions for soil and dust ingestion by children from 1 to 7 years obtained using bootstrapping (mg/d)

	Soil (outdoor and tracked-in)		Soil (outdoor)	Dust (indoor)	Soil and dust
	Distribution of arithmetic mean	Distribution of 95th percentile			
Arithmetic mean	63	97	39	45	87
Standard deviation	11	16			
95th percentile	81	124			

Table 6.8 Statistical descriptors of the distributions for soil ingestion for adults obtained using bootstrapping (mg/d)

Soil (outdoor and tracked-in)	Distribution of arithmetic mean	Distribution of 95th percentile
Arithmetic mean	46	81
Standard Deviation	9	13
95.0%	60	104

the 75th percentile is used for adults to estimate the upper bound soil ingestion rate (as proposed by Calabrese 2003) the values range between 37 and 120 mg/d with an arithmetic mean of 65 mg/d. When these data are subjected to bootstrapping, the distribution based on the arithmetic mean values is characterized by a mean soil ingestion rate of 46 mg/d (95% = 60 mg/d). The distribution around the 95th percentile has a mean value of 81 (95% = 104 mg/d; Table 6.8).

So far, most soil ingestion values discussed were derived from tracer studies. As mentioned in the introduction, tracer studies are just one of several possible approaches to estimate soil ingestion rates. Estimates of soil ingestion can equally well be derived from studies on hand-loading, provided sufficient details are available on micro activity patterns and transfer factors. The general equation used for estimating non-dietary soil and dust ingestion via hand-to-mouth contact involves the product of hand residue or soil loading, hand-to-mouth frequency, hand surface area mouthed per event, the efficiency of hand-to-mouth transfer and exposure duration. These parameters depend on the characteristics of the surfaces and contaminants involved and the activity the human being is involved in. All of them are fraught with large variability. Van Holderbeke et al. (2008) calculated soil and dust ingestion values based on data published on hand-loading by Holmes et al. (1999) for different target populations, e.g., day-care kids, children playing indoors, gardeners, archaeologists, et cetera. Estimated average dust ingestion rates for children (1–4 years) playing inside for 10 h/day amount to 50 mg/d, which is comparable to estimated dust ingestion values (45 mg/d) from tracer studies. Other dust ingestion rates for indoor playing children range from 5 to 20 mg/d. However, the age range considered was broad (3–13 years), and a higher ingestion rate is often assumed for the younger children. The estimated soil ingestion for gardeners, utility workers and archaeologists, assuming a working day of 8 hours, ranges from 20 to 55 mg/working day.

Although this methodology is suitable for calculating detailed time-dependent soil ingestion rates (hourly ingestion rates), it should be kept in mind that the calculations are based on micro activity data mainly published in non-European literature. More extensive data on, for example, time activity patterns for children in different European countries may be required in order to obtain more accurate data for local Risk Assessments. As well, age ranges are often broad and there still is a need for more data on separate indoor and outdoor activities for children.

Van Holderbeke et al. (2008) used the Biokinetic model comparison methodology based on the validation of exposure models for contaminants to estimate soil and dust ingestion rates. Analysis of the available data suggest that average soil/dust ingestion rates for children are not higher than 100 mg/d and may well be lower, down to 50 mg/d. However, the data do not provide a better estimate for a specific site than this more or less safe 100 mg/d, due to the uncertainties in soil and dust contribution and the bioavailability in soil and dust. The data also do not provide an estimate for the contribution of soil and dust, nor for the soil ingestion rates for adults.

Another approach used by Van Holderbeke et al. (2008) is trying to establish an empirical relation between contaminant concentrations in the environment and Biomonitoring data. From their analysis Van Holderbeke et al. (2008) conclude that typical soil ingestion rate values for children are below 50 mg/d, both for soil and dust. The obtained dust and soil ingestion values range from 20 to 70 mg/d. Insufficient data exist to estimate soil ingestion values for adults using this approach.

The majority of studies aims to estimate "normal" unintentional soil and dust ingestion by children and adults. However, short-term or longer-term ingestion of larger amounts of soil (g/day) can occur. Due to the limited data, it is difficult to quantify pica ingestion rates. High-end estimates from tracer studies or survey response information, however, can support the estimations.

6.2.4 Soil Ingestion Rates Recommended by International Regulatory Bodies

The available studies are used by several international regulatory bodies to recommend or impose soil and dust ingestion rates for use in Human Health Risk Assessments.

An overview of soil ingestion values recommended by international regulatory bodies and ingestion values used in different exposure models (Swartjes 2002), is given in Table 6.9. The values used in exposure models are often at the high end of the available ranges. This could be due to the fact that they are based on older studies, but as well on the fact that these models are often screening models and a high level of protection is envisaged. From the information available, it is not always clear whether the values relate to soil ingestion or to soil and dust ingestion combined. In most models, however, the value relates to soil ingestion, only.

6.2.5 Representativeness of Soil and Dust Ingestion Rates

The estimation of appropriate site-specific soil and dust ingestion rates remains difficult. Tracer element studies were mostly carried out over a relatively short time period, thus estimated intake rates may not reflect long-term patterns. In this context DEFRA (2006) comments "that current high-end exposures (derived from the short two-week studies) overpredict typical longer-term exposure because they fail to account for variability in the underlying data which when extrapolated to the longer term leads to 'regression to the mean' ".

Table 6.9 An overview of soil ingestion values recommended by international regulatory bodies and ingestion values used in different exposure models

	Values proposed by regulatory authorities											Values used in models						
		US_EPA 1997	US_EPA 2008	CEPA 1994	EDH 1993	CCME 1996	MOEE 1995	OEHHA 1997	DEP 2002	Japan 2001	Danish EPA 2005	Clea	CETOX human	CSOIL 2000	ROME 01	Vlier Humaan	Sweden	France
	Scenario																	
Children	CTE (Central Tendency Exposure)	90 (100)	30 (0.5 – <1 year) – soil 60 (0.5 – <1 year) – soil + dust	35 (0 – 0.5 year)	70 (1–7 year)	20 (0 – 0.5 year)	35 (0 – 0.5 year)	200	100	200	(200) 100	100	1000	100	200	50	150	150
		106 (soil + dust)	50 (1 – <21 year) – soil 100 (1 – <21 year) – soil + dust	50 (0.5–4 year)		80 (0.5–4 year)	80 (0.5–4 year)											
				35 (5–11 year)		20 (5–11 year)	80 (5–11 year)											
				20 (12–19 year)		20 (12–19 year)	20 (12–19 year)											
	RME (Reasonable Maximum Exposure)	236									400	300 (P95)						
		449 (soil + dust)																
Adults	CTE	50		20	20	20	20	100	50	100			25	50	100	20	50	50

Nearly all tracer element studies were conducted in the summer period, a time when soil contact probably will be more frequent than in other periods of the year. The majority of the soil and dust ingestion studies are of USA origin and assessment of their applicability outside the USA is limited.

In addition to inter-individual differences in behaviour, a number of regional factors can influence soil and dust ingestion rates, thus leading to significant variation within regions and even countries. Climate influences the time spent outside and the possibility for soil contact (e.g., snow cover). The difference between urban areas with almost no gardens and little green areas compared to rural areas where children have free access to uncovered areas will influence ingestion. Also, the "dustiness" of the playing environment will have a role in the potential for soil and dust ingestion. Often, contaminated regions are more dusty because of emission sources, therefore resulting in potentially higher ingestion rates.

The Biokinetic validation studies show that a certain set of soil/dust ingestion rates predicts intake well for one contaminated site, but not for another. This can be related to absolute ingestion rates, but also to differences in the contribution from the various soil and dust compartments (garden soil, community soil, house dust, school dust).

Although we can define both child-specific and regional factors that influence soil and dust ingestion rates, the quantification of their influence still is difficult. The results of the assessment of the scientific data constitute the basis for deriving default values or distributions, whereas consideration of region- or site-specific information could be used to modify default values.

6.3 Conclusions

Different methodologies for the derivation of soil and dust ingestion rates for use in Human Health Risk Assessment are available, but all have their advantages and disadvantages. Moreover, the collection of the information is time-consuming and requires active involvement of experimental participants. Therefore, the number of studies is limited and the number of participants per study is low, so that a decent statistical analysis is difficult.

A first requirement for analysis of the studies is a clear definition of what is meant by soil and dust. Then, the studies can be analyzed in the light of these definitions. Human Health Risk Assessment for contaminated sites requires at least the consideration of ingestion from outdoor soil and from soil-derived indoor dust. *Tracer methodology studies* probably provide the most adequate estimates of soil ingestion rates, but are limited to short-term estimates and do not allow us to distinguish between soil and dust. *Hand-loading studies* can be designed such that information is collected for the relevant micro-environments, but the interpretation requires assumptions about transfer parameters, as such, introducing substantial uncertainty. At the moment, *Biokinetic modelling studies* can only provide estimates of ranges, but are nevertheless useful as a complementary line of evidence.

Current analysis indicates that average soil ingestion rates for children are below 100 mg/d and should be around 50 mg/d. Average dust ingestion rates are within the same range, although children spend more time indoors than outdoors. This means that soil or dust ingestion per unit of time is higher outdoors than indoors, probably due to the more vigorous activities children are involved in outdoors and the "unlimited" amounts of soil available. Higher percentiles of soil ingestion are estimated to be around 100 mg/d. Estimates for pica-behaviour are highly uncertain, but are in the order of g/day.

Best ranges or distributions for soil and dust ingestion rates can be derived from the available studies and default values can be set. Region- or site-specific information on e.g. climate, urbanization, dustiness could be used to justify choices from these data.

References

ATSDR (2001) Summary report for the ATSDR soil pica workshop, June 2000, US Department of Health and Human Services, Atlanta, Georgia

AuYeung W, Canales R, Beamer P, Ferguson AC, Leckie JO (2004) Young children's mouthing behavior: An observational study via videotaping in a primarily outdoor residential setting. J Children's Health 2(3–4):271–295

Bierkens J, Cornelis C (2006) Opstellen van gebruiksafhankelijke waarden voor bodemingestie (No2006/IMS/R/387) (in Dutch). VITO, Mol, Belgium

Binder S, Sokal D, Maughan D (1986) Estimating soil ingestion: the use of tracer elements in estimating the amount of soil ingested by young children. Arch Environ Health 41(6):341–345

Calabrese EJ (2003) Letter from Dr. E. Calabrese to Kevin W. Holtzclaw, manager general electric company: soil ingestion rates. http://www.epa.gov/region1/ge/thesite/restofriver/reports/final_hhra/comments/generalelectric/AttachE.pdf. Accessed July 2003

Calabrese EJ, Pastides H, Barnes R, Edwards C, Kostecki PT (1989) How much soil do young children ingest: an epidemiologic study. In: Petroleum contaminated soils. Lewis, Chelsea, MI, pp 363–397

Calabrese EJ, Stanek EJ (1992) What proportion of household dust is derived from outdoor soil? J Soil Contamin 1(3):253–263

Calabrese EJ, Stanek EJ (1995) Resolving intertracer inconsistencies in soil ingestion estimation. Environ Health Perspect 103(5):454–456

Calabrese EJ, Stanek EJ, Barnes RM (1996) Methodology to estimate the amount and particle size of soil ingested by children: Implications for exposure assessment at waste sites. Regul Toxicol Pharmacol 24:264–268

Calabrese EJ, Stanek EJ, Gilbert CE, Barnes RM (1990) Preliminary adult soil ingestion estimates: results of a pilot study. Regul Toxicol Pharmacol 12:88–95

Calabrese EJ, Stanek EJ, Pekow P, Barnes RM (1997) Soil ingestion estimates for children residing on a Superfund site. Ecotoxicol Environ Safety 36:258–268

Camann DE, Majumdar TK, Geno PW (2000) Evaluation of saliva and artificial salivary fluids for removal of pesticide residues from human skin. Southwest research institute, San Antonio TX and ManTech Environmental Technology, Research Triangle Park, NC, for US EPA. National Exposure Research Laboratory, Research Triangle Park, NC. EPA/600/R00/041

CCME (1996) A protocol for the derivation of environmental and human health soil quality guidelines. Canadian council of ministers of environment. Winnipeg, MB, CCME-EPC-101E. En 108-4/8-1996E. ISBN 0-662-24344-7

CEPA (1994) Human health risk assessment for priority substances. Canadian environmental protection act. Health Canada. ISBN 0-662-22126-5

Clausing P, Brunekreef B, Van Wijnen JH (1987) A method for estimating soil ingestion by children. Int Arch Occup Environ Health (W. Germany) 59(1):73–82

Cohen Hubal EA, Suggs JC, Nishioka MG, Ivancic WA (2005) Characterizing residue transfer efficiencies using a fluorescent imaging technique. J Exposure Anal Environ Epidemiol 15(3):261–270

Cornelis C, Swartjes FA (2007) Ontwikkeling van een geharmoniseerde methodiek voor beoordeling van gezondheidsrisico's door bodemverontreiniging in de Kempenregio (in Dutch). VITO, Mol, Belgium and RIVM, Bilthoven, the Nederlands

Davis S, Mirick D (2006) Soil ingestion in children and adults in the same family. J Exposure Anal Environ Epidemiol 16:63–75

Davis S, Waller P, Buschbon R, Ballou J, White P (1990) Quantitative estimates of soil ingestion in normal children between the ages of 2 and 7 years: population based estimates using aluminium, silicon, and titanium as soil tracer elements. Arch Environ Health 45:112–122

DEFRA (2006) Assessing risks from land contamination – a proportionate approach. Soil guideline values: the way forward. http://www.defra.gov.uk/environment/land/contaminated/pdf/clan6–06.pdf, November 2006

DEP (2002) Massachusetts department of environmental protection. Calculation of an enhanced soil ingestion rate, updates: appendix B, guidance for disposal site risk characterization – in support of the Massachusetts contingency plan (1995)

Egeghy PP, Sheldon LS, Fortmann RC, Stout DM II, Tulve NS, Cohen Hubal E, Melnyk LJ, Morgan M, Jones PA, Whitaker DA, Croghan CW, Coan A (2007) Important exposure factors for children. An analysis of laboratory and observational field data characterizing cumulative exposure to Pesticides. NERL, National Environmental Research Institute

Finley BL, Scott PK, Mayhall DA (1994) Development of a standard soil-to-skin adherence probability density function for use in monte carlo analyses of dermal exposure. Risk Anal 14 (4):555–569

Freeman CG, Jimenez M, Reed KJ, Gurunathan S, Edwards RD, Roy A, Adgate JL, Pellizzari ED, Quackenboss J, Sexton K, Lioy PJ (2001) Quantitative analysis of children's microactivity patterns: the Minnesota children's pesticide exposure study. JEAEE 11:501–509

Hager JM (2005) Determination of the external soil fraction (ESL) and the outdoor soil lead contribution (OSLC) for Syracuse, NY. State University of New York, Syracuse, NY

Holmes KK Jr, Shirai JH, Richter KY, Kissel JC (1999) Field measurement of dermal soil loadings in occupational and recreational activities. Environ Res Section A 80:148–157

Kissel JC, Shirai JH, Richter KY, Fenske RA (1998) Empirical investigation of hand-to-mouth transfer of soil. Bull Environ Contamin Toxicil 60:379–386

Kuusisto S, Lindroos O, Rantio T, Priha E, Tuhkanen T (2007) PCB contaminated dust on indoor surfaces – Health risks and acceptable surface concentrations in residential and occupational settings. Chemosphere 67:1194–1201

MOEE (1995) Health risk assessment of mercury contamination in the vicinity of ICI forest products Cornwall, Ontario. Ontario ministry of environment and energy. May 1995. PIBS 3352

OEHHA Office of Environmental Health Hazard Assessment, California (2007) Safe drinking water and toxic enforcement act of 1986 (proposition 65) notice to interested parties June 8, 2007 notice of a public workshop for comments on a draft interpretive guideline for hand-to-mouth transfer of lead exposure from fishing tackle products

Oomen AG, Lijzen JPA (2004) Relevancy of human exposure via house dust to the contaminants lead and asbestos. RIVM Report 711701037/2004. RIVM, Bilthoven, the Netherlands

Rutz E, Valentine J, Eckart R, Yu A (1997) Pilot study to determine levels of contamination in indoor dust resulting from contamination of soils. J Soil Contamin 6(5):525–536

Sedman RM, Mahmood RJ (1994) Soil ingestion by children and adults reconsidered using the results of recent tracer studies. Air waste 44(2):141–144

Stanek EJ, Calabrese E, Barnes R, Pekow P (1997) Soil ingestion in adults—results of a second pilot study. Ecotoxicol Environ Safety 36(3):249–257

Stanek EJ, Calabrese EJ (1995a) Daily estimates of soil ingestion in children. Environ Health Perspect 103(3):276–285

Stanek EJ, Calabrese EJ (1995b) Soil ingestion estimates for use in site evaluations based on the best tracer method. Human Ecol Risk Assess 1:133–156

Stanek EJ, Calabrese EJ (2000) Daily soil ingestion estimates for children at a superfund site. Risk Anal 20(5):627–635

Stanek EJ, Calabrese EJ, Barnes R (1999) Soil ingestion estimates for children in Anaconda using trace element concentrations in different particle size fractions. Human Ecol Risk Assess 5(3):547–558

Stanek EJ, Calabrese EJ, Zorn M (2001a) Biasing factors for simple soil ingestion estimates in mass balance studies of soil ingestion. Human Ecol Risk Assess 7(2):329–355

Stanek EJ, Calabrese EJ, Zorn M (2001b) Soil ingestion distributions for Monte Carlo risk assessment in children. Human Ecol Risk Assess 7(2):357–368

Swartjes F (2002) Variation in calculated human exposure comparison of calculations with seven European human exposure models. RIVM report 711701030 /2002. RIVM, Bilthoven, the Netherlands.

Thompson KM, Burmaster DE (1991) Parametric distributions for soil ingestion by children. Risk Anal 11:339–342

Trownbridge PR, Burmaster DE (1997) A parametric distribution for the fraction of outdoor soil in indoor dust. J Soil Contamin 6(2):161–168

Tulve NS, Suggs JC, McCurdy T, Cohen Hubal EA, Moya J (2002) Frequency of mouthing behavior in young children. JEAEE 12:259–264

US EPA (1997) Exposure factors handbook. EPA/600/P-95/002F.US-EPA

US EPA (1999) IEUBK model soil/dust ingestion rates – guidance document. O.o.E.R.R. response, United States environmental protection agency

US EPA (2001) Policy paper from the science advisory council for exposure, policy number 12. Recommended revisions to the standard operating procedures (SOPs) for residential exposure assessments, Washington, DC

US EPA (2006) Child-specific exposure factors handbook (external review draft), September 2006, EPA/600/R/06/096A

US EPA (2008) Child-specific exposure factors handbook (2008) US environmental protection agency, Washington, DC, EPA/600/R-06/096F, 2008

US EPA OoSWaER (1998) Short sheet: IEUBK model mass fraction of soil in indoor dust (MSD) variable. Report no: #540-F-00–008, US EPA, Washington DC

Van Holderbeke M, Cornelis C, Bierkens J, Torfs R (2008) Review of the soil ingestion pathway in human exposure assessment. Study in support of the BeNeKempen project – subgroup on harmonization of the human health risk assessment methodology. http://www.ovam.be/jahia/Jahia/cache/offonce/pid/176?actionReq=actionPubDetail&fileItem=2026. Accessed June 2008

Van Wijnen JH, Clausing P, Brunekreff B (1990) Estimated soil ingestion by children. Environ Res 51:147–162

Xue J, Zartarian V, Moya J, Freeman N, Beamer P, Black K, Tulve N, Shalat S (2007) A meta-analysis of children's hand-to-mouth frequency data for estimating nondietary ingestion exposure. Risk Anal 27(2):411–420

Zartarian VG, Xue J, Ozkaynak HA, Dang W, Glen G, Smith L, Stallings C (2005) A probabilistic exposure assessment for children who contact CCA-treated playsets and decks using the stochastic human exposure and dose simulation model for the wood preservative scenario (SHEDS-WOOD), Final report. US Environmental Protection Agency, Washington, DC, EPA/600/X-05/009

Chapter 7
Oral Bioavailability

Mark R. Cave, Joanna Wragg, Sébastien Denys, Catherine Jondreville, and Cyril Feidt

Abstract Soil ingestion is a key exposure pathway in Human Health Risk Assessment for contaminants in soil. The theory and mechanisms of how contaminants in a soil enter the human body through the gastrointestinal tract are outlined. The methods available for measuring human exposure using human, animal and validated in-vitro laboratory methods are described and contrasted. The role of the physico-chemical properties of the soils that control the bioavailability of contaminants are summarised. Finally, examples of how bioavailability/bioaccessibility studies of soils from both anthropogenic and geogenic origin are discussed along with the criteria required for deciding whether bioavailability data should be used in a Human Health Risk Assessment.

Contents

M.R. Cave (✉)
British Geological Survey, Keyworth, Nottingham, UK
e-mail: mrca@bgs.ac.uk

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_7,

7.1 Theory of Availability

Soil ingestion (see Chapter 6 by Bierkens et al., this book; Pausten bach 2000) is a key exposure pathway in Human Health Risk Assessment (HHRA) for contaminants in soil (see Chapter 5 by Swartjes and Cornelis, this book). To date, research regarding the human bioavailability of soil contaminants, via the ingestion exposure pathway, has concentrated on inorganic contaminants, and therefore this chapter is focussed on these contaminants. The priority inorganic contaminants have been arsenic (Beak et al. 2006; Bruce et al. 2007; Davis et al. 1996a; Ellickson et al. 2001; Juhasz et al. 2008; Koch et al. 2005; Palumbo-Roe and Klinck 2007; Laird et al. 2007; Roberts et al. 2007b; Wragg et al. 2007) and lead (Abrahams et al. 2006; Denys et al. 2007; Drexler and Brattin 2007; Hamel et al. 1999; Oomen et al. 2002; Ren et al. 2006; Ruby et al. 1993, 1999; Van de Wiele et al. 2005, 2007), but work has been carried out on other metals such as cadmium (Tang et al. 2006), chromium (Fendorf et al. 2004; Nico et al. 2006; Pouschat and Zagury 2006; Stewart et al. 2003a, b), nickel (Barth et al. 2007; Hamel et al. 1998; Hansen et al. 2007) and mercury (Cabanero et al. 2007; Davis et al. 1997). As such, the inorganic contaminants referred to in this chapter include metals and metalloids such as arsenic, lead, cadmium, chromium and nickel, but not contaminants such as cyanide. For the organic contaminants, recent insights and projects have concentrated on polyaromatic hydrocarbons (PAHs) (Gron 2005; Tilston et al. 2008) and dioxin and furans (Wiettsiepe et al. 2001). However, for the degradable contaminants, the estimation of the bioavailable fraction appears to be more complicated than for inorganic contaminants, as the metabolites and degradation products of these contaminants have to be taken into consideration. For example, Van de Wiele et al. (2005) showed that colon microbiota plays an important role in the biodegradation of PAH, thus influencing the outcomes of the bioaccessibility measurement undertaken.

In Risk Assessments using exposure models, it is often assumed that inorganic contaminants ingested via soil are absorbed to the same extent as the available form that was used during the toxicological assessment. Thus, assuming that the rate of absorption of ingested metal is independent of the matrix in which it is included and of its chemical form (speciation). However, the binding of inorganic contaminants to solid phases in soils may render them unavailable for absorption; similar mechanisms have been extensively shown to modulate uptake of metals by plants and metal speciation plays an important part in the uptake process (see Chapter 8 by McLaughlin et al., this book).

It can then be assumed that only a fraction of the inorganic contaminant will be absorbed by the receptor (human or animal), this fraction is called the *bioavailable fraction*. As oral exposure is being considered the term oral bioavailability is used. When using a Risk Assessment model, it is therefore very important to understand whether the underlying algorithm incorporates bioavailability in the oral exposure pathway. This chapter focuses on oral bioavailability (OB) to humans.

7.1.1 Oral Bioavailability

Bioavailability can be defined as the fraction of an ingested contaminant that is absorbed and reaches the systemic circulation where it may then cause adverse effects on human health. Oral bioavailability can be assessed by comparing the internal doses obtained after oral administration to that of intravenous administration of the contaminant. The absolute oral availability is defined as the ratio of the oral administration (AUC_{PO}) to the intravenous administration (AUC_{IV}). The internal dose is often related to the area under the curve (AUC), using the relationship between time and plasmatic concentrations of the contaminant (Fig. 7.1). However, other end-points and methods of determining the oral bioavailability of contaminants are commonly used and briefly discussed in Section 7.1.2.

Oral bioavailability is of toxicological interest because the possible adverse effects, caused by the contaminant, on the exposed human subject are related to the internal dose. In order to depict the effect of soil properties on contaminant availability, this concept can be divided into several steps; *accessibility*, *absorption* and *metabolism*.

The most commonly used definition of oral bioavailability is based upon a two-step model with three compartments shown in Fig. 7.2 and Eq. (7.1):

$$F_{\mathrm{Bacc}} \times F_{\mathrm{Abs}} = F_{\mathrm{Bava}} \tag{7.1}$$

A slightly different definition (used by the Dutch National institute of public health and the environment, *RIVM*) includes a third step with liver action (metabolism and possibly bile excretion) on the contaminant as defined by Eq. (7.2):

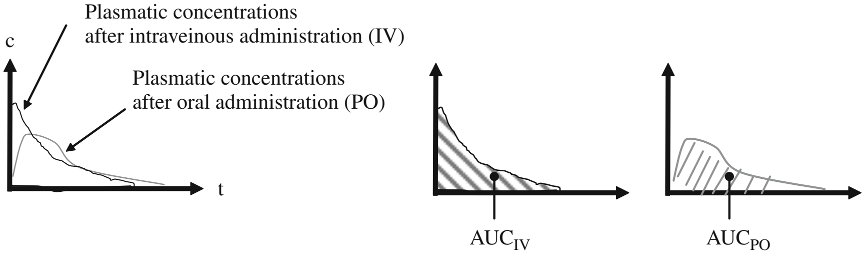


Fig. 7.1 Oral Bioavailability pharmako-kinetics definition

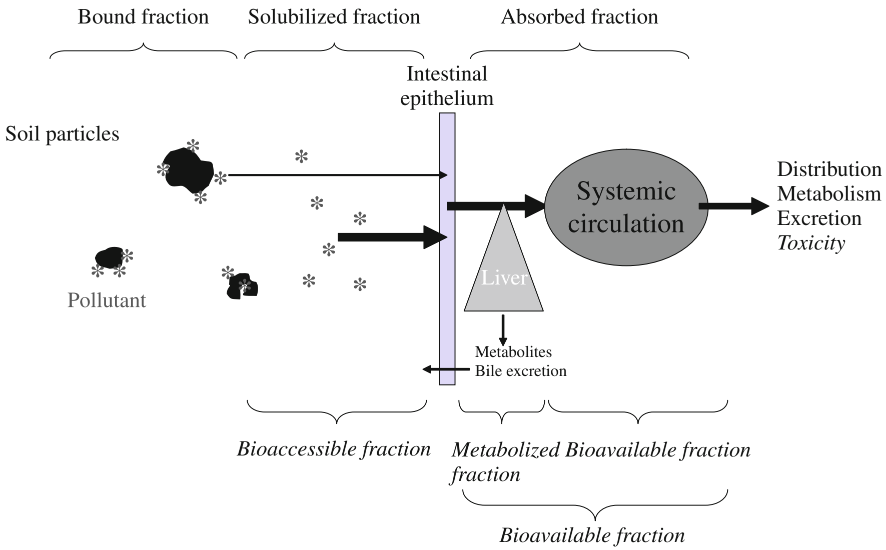


Fig. 7.2 Steps involved in oral bioavailability

$$F_{\mathrm{Bacc}} \times F_{\mathrm{Abs}} \times Fn_{\mathrm{Met}} = F_{\mathrm{Bava}} \tag{7.2}$$

where:

F_{Bava} is the bioavailable fraction;
F_{Bacc} is the fraction of ingested contaminant that is bioaccessible;
F_{Abs} is the fraction of accessible contaminant that is absorbed;
Fn_{Met} is the fraction of absorbed contaminant that is not metabolised in liver (= 1–fraction metabolised).

7.1.1.1 Accessibility

Contaminants that are released from soil particles within the gastrointestinal tract are considered to be *bioaccessible*. For a contaminant in soil to be bioaccessible to a human via ingestion, it must be released from the soil into solution in the gastrointestinal tract in a form that can be absorbed into the blood stream. Conditions in the stomach, such as pH and residence time, may vary, depending on whether the individual is in a fed or fasted state. It is generally considered, when simulating human gastro-intestinal conditions, that a fasted state is likely to mobilise the highest amount of metals (Maddaloni et al. 1998; Van de Wiele et al. 2007). The fasted state is, therefore, considered to be the most conservative for inorganics. Conversely, for organic contaminants, the opposite is considered to be conservative, i.e. the fed state, for the reasons discussed in Section 7.1.1.2.

The main factors that control the release of contaminants are discussed in the following sections. The amount of contaminant liberated from the soil is considered to be the maximum amount that is potentially available for absorption. This is called *the bioaccessible fraction*. The accessible fraction is produced mainly in the upper part of the digestive tract, particularly in the stomach compartment, in the case of metals. The extent to which ingested contaminants are rendered accessible depends of physicochemical conditions in the digestive tract, as well as of transit time. Both are modulated by the presence of food and soil in the digestive tract. The differences between fed and fasted states are:

- The presence or absence of food;
- The amount of enzymes such as bile in the intestinal fluid (these can be very different in the fed and fasted states); and
- The pH conditions under which the model is employed (when food is present the stomach pH is much higher).

Under fed conditions, the model should include an estimate of the nutritional status of the receptor under investigation (for the RIVM fed state model, baby food is included), an increase in the stomach pH and an alteration in the amount of enzymes included in the test system (some are increased, such as the amount of bile added, but others decreased) (Versantvoort et al. 2004).

7.1.1.2 Absorption

Absorption occurs predominantly in the small intestine. This is a highly developed organ with a large surface area specifically for this purpose. The walls of the small intestine comprise of finger-like projections, of about one millimetre long, called villi, which are in turn covered by numerous micro-villi as shown in Fig. 7.3. The villus contain blood and lymphatic capillaries, which transport nutrients, and for the purposes of this discussion, contaminants to the rest of the human body. Enterocytes (absorptive cells, which are predominantly responsible for absorption

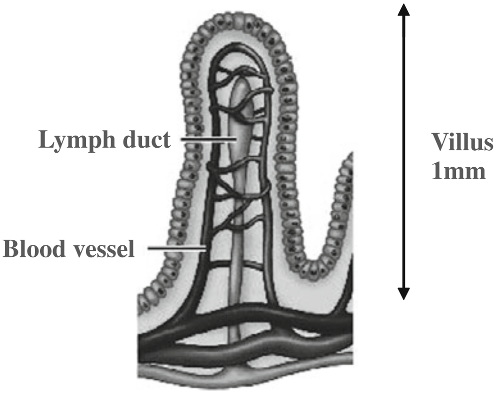


Fig. 7.3 Intestinal villi, finger-like projections of about 1 mm long

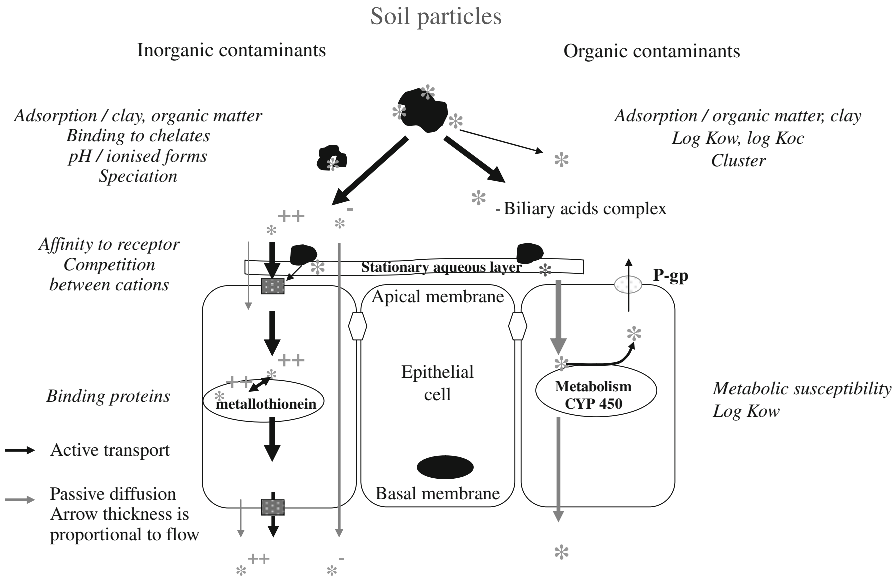


Fig. 7.4 Mechanisms of absorption

(Oomen 2000)) in conjunction with a heterogeneous mass of other cell types such as goblet cells (those that secrete mucin) and endocrine cells populate the surface of each villi.

Because of the different properties of nutrients/contaminants, absorption can occur either by the para or trans cellular route, as shown in Fig. 7.4. Figure 7.4 shows that the epithelial cells bound together at their luminal facing ends by specialized facings known as 'zonnulae occludens', which carry out a number of functions such as the preservation of the cell sheet integrity and the formation of a boundary between apical surfaces and basolateral membranes. For the transcellular route, movement occurs across both apical and basolateral membranes of the intestinal lumen, however for the para cellular route movement of nutrients/contaminants occurs by circumventing these membranes. The small spaces between the cells mean that paracellular absorption can only occur for the transport of small hydrophobic molecules. Passive forces drive this process, with the rate limiting forces which determine the extent or magnitude of transport determined by the permeability between the cells. Larger molecules are transported by the transcellular route, which involves the movement across both the apical and basolateral membranes. The accepted model for transcellular transport assumes that substances absorbed by this route enter the cell across the apical and leave via the basolateral membrane and that the opposite is true for secretions. The mechanisms of transcellular transport may occur by:

- Transport by a specific carrier, either active or facilitated. Active transport is the mechanism by which proteins (after transformation into amino acids or small

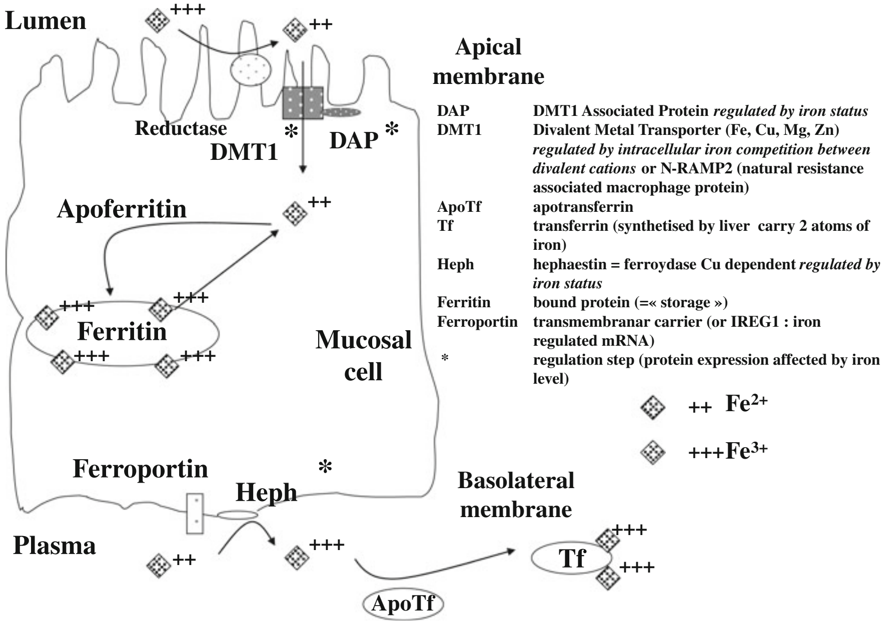


Fig. 7.5 Intestinal absorption of iron and its regulation

peptides), carbohydrates and ions are absorbed. This transport route uses carrier molecules to ferry nutrients, or accessible contaminants, across the intestinal membrane. In addition to carrier molecules, this form of transport requires energy in the form of adenosine triphosphate (ATP). However, some carbohydrates, such as fructose, require special carrier molecules for transport, but not the addition of energy via ATP. This mechanism of transport is known as facilitated.

- Passive diffusion, where neither energy or special carrier molecules are required. Examples include organic contaminants that interact with the bile salts from the liver and form mixed micelles, which are able to diffuse freely through the mucosal cell wall.
- Transcytosis or pinocytosis, where ‘a small volume of the intestinal fluid is invaginated by the mucosal cell wall to form an endocytotic vesicle’ (Oomen 2000).

Figure 7.5 illustrates the process of the absorption of contaminants, specifically iron, through the mucosal cell wall (Amshead et al. 1985; Bridges and Zalups 2005; Kim et al. 2007; Park et al. 2002; Sharp 2003; Tallkvist et al. 2001). For human/animal receptors in a state of low dietary mineral intake (i.e. mineral deficient), the primary routes of absorption of bioaccessible cations occurs via the active transport mechanism using transmembranar proteins as the carrier molecules. Carrier molecules include the Divalent Metal Transporter 1, DMT1 and the Zinc

Transporter 4 Precursor, ZIP4 (not presented in Fig. 7.5). On ingestion of excessive doses of mineral salts, the gradient between the intestinal lumen and plasma or cell will be so large that it allows for passive diffusion (proportional to the concentration between the two compartments) to occur. In this instance, the process of active transport from the lumen to systemic circulation is decreased, because the absorption of mineral salts would be too high, and there is an increase in the excretion of the mineral salts e.g. through sweating in order to protect the organism from poisoning.

Divalent cations, such as calcium, cadmium, copper, magnesium, manganese, iron and zinc, exhibit similar common behaviour under the acidic pH conditions in the very upper intestine (*proximal duodenum*):

- ionisation and solubilisation in the intestinal fluids;
- binding to transmembranar proteins;
- active process (energy consumption) during apical crossing;
- intracellular release as ions and chelation.

For example, iron is absorbed in the proximal duodenum. In order for the absorption process to be efficient, an acidic environment is required (where the solubilisation and ionisation processes occur), however, these conditions may be hindered by the ingestion of antacids etc., that interfere with gastric acid secretion. Ferrireductase reduces ferric to ferrous iron in the duodenal lumen and after binding to the transmembranar protein DMT-1, the iron is cotransported with a proton into the absorbtive intestinal cells. The transmembranar protein DMT-1 is not iron specific, it is a transporter of many divalent metal ions. Once inside the intestinal cell, the absorbed iron may follow one of two major pathways, which is dependant on the dietary status of the host and the iron loading already present in the cell. For iron abundant states, the iron within the cell is trapped by incorporation into ferritin. This iron is not transported to the blood and the iron is lost when the cell dies. Whereas under conditions where there is a paucity of iron; the absorbed iron is exported from the cell via the transporter ferroportin, which is found in the basolateral membrane, and transported through the body after binding to the iron-carrier transferrin (i.e. intracellular release and chelation).

During the protein binding steps in the active transport process cations may interact, as antagonists (i.e. oppose the action of additional cations or bind to the receptor without producing a response). Thus the absorption of specific bioaccessible cations to the binding/carrying proteins may be affected by the presence of other cations in the target organ and competition with other bioaccessible contaminants. For example, this regulation means that cadmium absorption can be affected not only by a competition with other ions to bind to the ligand site, but also by the iron status of the receptor.

For some metals bound to amino-acids in soil, absorption can occur via amino acid carriers. In the case of copper, the amino acid carrier is histidine, whereas for selenium, methionine complexes are formed. After their translocation to plasma, metals can be bound to specific or non-specific carriers. Specific carriers include transferrin (a blood plasma protein for iron delivery), transcobalamin (a group of proteins (of intestinal cells) that bind to cyanocobalamin (vitamin B12) and transport it to other tissues) and nickeloplasmin (a nickel containing protein), the non-specific class of carriers includes albumin and amino-acids.

When considering the organic contaminants of major concern, i.e. dioxins, polychlorobiphenyls (PCBs) and polyaromatic hydrocarbons (PAHs), absorption through the intestinal epithelium is commonly described by a passive diffusion model (Cavret and Feidt 2005). However, because there is a large variation in behaviour amongst the organic contaminants of interest, intestinal absorption of organic contaminants will not be covered in great detail in this chapter. In the passive diffusion model, the diffusion flow is proportional to the concentration gradient between the luminal and plasmatic compartments, according to their lipid contents. Most organic contaminants of interest are not readily soluble in aqueous digestive fluid, because of their hydrophobicity, and are conveyed by micellar solutions of conjugated bile salts. Bile acids are facial amphipathic molecules, i.e., they contain both hydrophobic (lipid soluble) and polar (hydrophilic) faces. Therefore, bile salts act as lipid carriers and are able to solubilise many organic contaminants by forming micelles, aggregates of lipids such as fatty acids, cholesterol and monoglycerides, which remain suspended in water and can reach the epithelium. Once the organic contaminants reach the double-layered membrane, their hydrophobicity favours their uptake across the epithelium (Fig. 7.3).

7.1.1.3 Metabolisation in the Liver

After absorption, contaminants migrate through the portal vein into the liver, in which they may undergo *biotransformation*. The most common transformation mechanisms are hydroxylation for lipophilic organic contaminants (PAHs, PCBs, dioxins), whereas the cationic metals are bound by proteins (Webb 1986). Arsenic, mostly under anionic form, may undergo methylation (Tseng 2009). Once absorbed and passed through the portal vein, contaminants may be:

- sequestrated within the liver (metals bound to proteins or PCDF bound to cytochrome P450);
- excreted in bile, either unaltered or in a metabolised/bound form; or
- released into the systemic circulation.

These processes are complex, because of the enzymes involved in the biotransformation process and the inducible production of binding proteins. The amount of contaminant metabolisation, via bile excretion, is dose dependant. At a low dose, the amount of activity will not be sufficiently efficient; however, continual exposure increases enzymatic activity due to DNA regulation, increasing the metabolism capacity followed by the rate of metabolism. The metabolism of a particular substance can be affected by both interspecies and inter-individual differences, thus contaminants that may be benign to humans may be highly toxic to other species and vice versa (Fowles et al. 2005).

The three-step approach described above, with accessibility, absorption and metabolisation is suitable for a pharmaceutical approach, i.e. relating to a drug which reaches its target. However, it may fail to properly represent the toxic potential of some contaminants, such as such as PAHs, which express their toxic (carcinogenic) effects through their hydroxy-metabolites. In this case the metabolism activates toxicity. This is an important consideration that needs to be addressed when considering the choice of method used to assess oral bioavailability. Because of analytical limitations, PAH metabolites such as hydroxy-PAH conjugated to sulphates, glucuronic acid or glutathione cannot be detected in plasma. Metabolites are thus excluded from the AUC determination, although they are the very origin of PAHs toxicity. Consequently, due to the biotransformation processes, the bioavailable fraction is underestimated when analysing only the native PAHs in blood (scenario 1). In addition, under circumstances when the analysis of PAH metabolites in plasma *is* technically feasible, the fraction that has been excreted via the bile is not taken into account in the bioavailability calculation (scenario 2). This results in biotransformation processes reducing the calculated bioavailable fraction, which is used in the resulting Risk Assessment, but not for the same reason and not to the same extent: in the first case 100% of the metabolites are ‘forgotten’ due to the analytical procedure chosen, in the second a proportion of the metabolites is lost through bile excretion. This example shows that term definition (the definition of what is considered to be bioavailable) and method of assessment are not truly independent, and that it is very difficult to obtain a *consensual absolute bioavailability.*

The methylation process is the primary process by which inorganic arsenic is metabolised in the body (Fowles et al. 2005). This process is generally considered to be a detoxification process, whereby the majority of the ingested arsenic is metabolised to form methyl arsenic species (mainly dimethylarsenic acid), which are then excreted in the urine. Analysis of urine has identified arsenate, arsenite, monomethylarsonic acid and dimethylarsinic acid (Cullen 2008) amongst other species, which have a lower affinity for tissue sulfhydryl groups than other inorganic arsenic species (Fowles et al. 2005).

7.1.2 Relative Bioavailability Factor

Since the bioavailability of inorganic contaminants when measured in a water soluble form is different when compared to the contaminants bound to soil, it is necessary to use a correction factor which takes the effect of the sample matrix e.g. a soil or an aqueous solution into account. This correction factor is considered to be the *relative oral bioavailability factor*. This factor is generally less than one, because the accessibility of contaminants in the soil is less than in a water soluble form through physico-chemical interactions with the solid phase of the soil.

The relative oral bioavailability factor of the contaminant in a specific matrix, e.g. a specific soil sample, is expressed as a percentage of the contaminant in the reference matrix, e.g. reference sample according to Eq. (7.3):

$$(\mathrm{AUC}_{\mathrm{PO}})_{\mathrm{TM}} \,/\, (\mathrm{AUC}_{\mathrm{PO}})_{\mathrm{RM}} \times 100 \tag{7.3}$$

where:

$(AUC_{PO})_{RM}$ is the area under the curve obtained from the oral administration of the contaminant in the reference matrix (RM), e.g. water;

$(AUC_{PO})_{TM}$ is the area under the curve obtained from the administration of the contaminant in the test matrix (TM), e.g. soil material.

This ratio is often between 0 and 1 (or 0 and 100%), because the reference material is chosen according to its physico-chemical properties in order to maximise absorption, i.e. water with ionisable metal forms or oil for organic contaminants.

A number of in-vivo animal trials (i.e. tests performed in a living organism) have been used to assess the bioavailability of soil contaminants, using a variety of species such as the monkey, pig, rat and rabbit (Drexler and Brattin 2007; Ellickson et al. 2001; Freeman et al. 1992, 1993, 1994, 1995, 1996; Roberts et al. 2007b; Schroder et al. 2004a). Unfortunately, rats and rabbits exhibit large differences in their digestive physiology compared to humans, rats also have a pre-stomach compartment with a very specific physiology. Although these models have a long history of use in toxicological studies etc., the aim for bioaccessibility and bioavailability studies with respect to contaminated land is to mimic human physiology and model the interactions occurring between the soil contaminants and the digestive fluids. In particular the interactions between digestive fluids and the soil are an important parameter to consider, the pH of rat digestive fluids is different to humans, and the practice of coprophagy (rats) and caecotrophy (rabbits), which the habit of feeding on excrement, may re-introduce the contaminants of interest to the gut. Although the soluble forms of the contaminants present may not be affected by the second passage through the GI tract, any interaction with organic matter and microbial communities may differ compared to the contaminant bound to the soil. An animal model considered to be physiologically similar to humans is the primate, however, few experiments have been conducted with this model (Roberts et al. 2007a). The reason for this is that this model is expensive and ethically fragile. The juvenile swine is considered to be a useful anatomical proxy for the human neonatal digestive tract (Miller and Ullrey 1987; Moughan et al. 1992) and has been successfully used to assess the bioavailability of lead (Casteel et al. 1996, 2006). Further considerations that determine the choice of animal model include the developmental speed, intestinal tract aging, the ratio between the bone and body mass etc., to this end, although rats and rabbits are adequate models, the juvenile swine model is a preferred candidate (Moughan et al. 1994; Rowan et al. 1994). Animal studies are expensive, criticized due to animal welfare and cannot be conducted with a large enough number and variety of contaminated soils. In-vitro tests (i.e. tests performed in a laboratory dish or test tube; an artificial environment (Latin for *in glass*)) allow researchers to overcome limitations of animal tests (Oomen et al. 2003; Ruby 2004; Ruby et al. 1999). However, animal testing still remains a necessary means to validate in-vitro methods (Schroder et al. 2004b).

Despite the problems associated with animal testing there are a number of different end-points that can be used to assess relative oral bioavailability. Because of

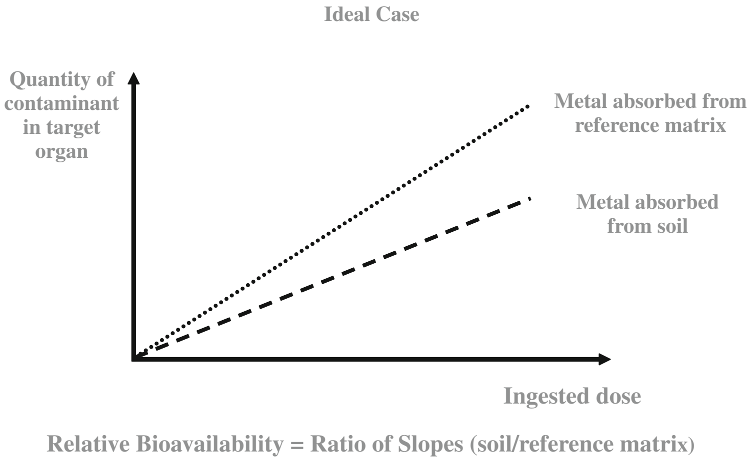


Fig. 7.6 Relative bioavailability value (RBV) assessment using dose response slopes

the complex animal manipulations required to obtain the plasmatic kinetics within juvenile swines, concentrations in alternative targets can be chosen. For lead, liver, kidney or bone concentrations at the end of the exposure period are considered to be proportional to the bioavailable fraction (Jondreville and Revy 2003). Unlike the area under the plasma lead concentration-time curve, these concentrations in target organs cannot be used to assess the actual amount which has been absorbed and has reached the systemic circulation. However, these concentrations do allow for a comparison to be made with the effects of the soil matrix on the bioavailability. In this case, the relative bioavailability can be expressed according to Eq. (7.4) and Fig. 7.6:

$$\text{RBA} = \text{Ratio of slopes (test matrix/reference matrix)} \tag{7.4}$$

Specific conditions are necessary to validate this calculation (Littell et al. 1997):

- there should be a linearity of response of both the reference and the test matrix over the dose range under investigation;
- the two lines should have a common intercept;
- the common intercept should be equal to the mean of the reference blank.

Depending on the contaminants of interest, specific target organs are relevant. The liver, kidney, bones and also urine give satisfactory responses for cations (lead, cadmium), urine and liver for anions (arsenic, antimony). In addition to using the AUC to study the oral bioavailability of contaminants, three additional primary methods are routinely employed (Kelley et al. 2002). Where the contaminant of interest is rapidly excreted, urine is a common end-point used in bioavailability studies as this fraction provides an estimate of the absorbed dose. Tissue concentration comparison, after administration of different forms of a contaminant, is useful for contaminants that preferentially accumulate in specific tissues. This type of end-point measurement provides an estimate of relative bioavailability. Additionally, the fraction of the dose excreted in the faeces may be measured; however this method reflects the amount of contaminant/dose that is not absorbed and the actual dose is calculated by subtracting the excreted amount from the initial amount dosed. As such, this method is subject to underestimating absorption if the contaminant is absorbed and subsequently excreted via bile back into the GI tract. For the organic contaminants and their metabolites, the database of available results is smaller compared to the metals and metalloids. Furthermore the choice of target organs/matrices for the metabolites is dependant on the metabolism process experienced by the individual contaminants. The fate of highly metabolised contaminants may be assessed by analysing their metabolites in urine, whereas unaltered contaminants may accumulate in adipose tissue.

7.1.3 Validation of Bioaccessibility Tests

In order to use a bioaccessibility test it is necessary to be able to demonstrate a mathematical, fit-for-purpose, relationship between the bioaccessible concentration, relative to the bioaccessibility of a soluble salt of the contaminant, in the test soil (as measured by an in-vitro test) and the bioavailable concentration, relative to the bioavailability of a soluble salt of the contaminant in the soil (as measured by an in-vivo study), see Section 7.1.2. In order to validate the in-vitro bioaccessibility test with the bioavailability data it is important that comparable units of measurement are used. The bioavailability of the element in the soil is almost always measured relative a water soluble salt of the element. It is therefore important that the comparison is made with the bioaccessibility of the soil relative to the same soluble salt. This may seem to be contradictory since, if the salt is completely soluble, the relative bioaccessibility will be the same as the absolute bioaccessibility. However, the water soluble salts are not necessarily fully soluble in the simulated stomach and intestine fluids particularly if the solubility of the element is reduced at the higher pH values found in intestinal fluid phase. Under these conditions, if relative bioaccessibility is not used, the bioaccessibility will always under predict the true relative bioavailability. For the test to be of pragmatic use, there are two major assumptions:

- the soils used in the validation test are representative of the soils that will be tested by the bioaccessibility test;
- the soluble salt used to calculate the relative bioavailabilities and the relative bioaccessibilities have the same biological action as the salt used in the toxicity test.

Previous studies (Basta et al. 2007; Drexler and Brattin 2007; Juhasz et al. 2007; Rodriguez et al. 1999, 2003; Schroder et al. 2003, 2004a) have shown that within the uncertainties on the measurement there is a linear relationship between bioaccessibility and bioavailability (or its relative counterpart). Monte-Carlo simulations show that for bioaccessibility to be predictive of bioavailability the straight line relationship must meet the following criteria:

- the soils used in the validation exercise must cover greater than 70% of the 0–100% of the bioavailability/bioaccessibility range;
- the between laboratory reproducibility of the bioaccessibility and bioavailability measurements must have a relative standard deviation of less than 20%;
- the r square value of the straight line should be greater than 0.7.

7.2 Influence of Soil Properties on Oral Bioaccessibility

The bioavailability of any contaminant bound to the soil depends upon the soil type, properties of the soil, the contaminant and the manner by which the contaminant has entered the soil (Selinus 2005).

7.2.1 pH

Soils generally have pH values (measured in water) from 4 to 8.5, due to buffering by aluminium at the lower end and calcium carbonate ($CaCO_3$) at the upper end of the range. In general, most divalent cationic forms of contaminants are less strongly absorbed, and therefore more bioaccessible, in acidic soils than they are at neutral or alkaline soils as demonstrated for lead by Yang et al. (2003) and for cadmium by Tang et al. (2006). In a study testing five Chinese sites (Tang et al. 2007), the bioaccessibility of both spiked and endogenous arsenic increased with increasing pH. A similar effect was observed by Yang et al. (2005), based on a study of 36 US soils. pH also affects other parameters such as the solubility of organic carbon, and the sorptive capacity of iron oxides and aluminium oxides and clays which are discussed in the next sections.

7.2.2 Soil Organic Matter

The soil organic matter content can vary widely among soils, from <1% to >70%. The organic matter is divided into non-humic and humic. The former consists of unaltered biochemicals which have not been degraded since they entered the soil, through production by living organisms. The latter are formed by secondary synthesis reactions involving micro-organisms.

The mechanisms that control the bioaccessibility of contaminants in the presence of soil organic matter have been described as (Selinus 2005):

- the adsorption of cations on negatively charged sites (ion exchange);
- the mobility and protection of some metal ions from adsorption through the formation of chelates with low molecular weight; and
- the retention of many contaminants in the higher molecular weight solid forms of humus.

Contaminants showing particularly high affinities to soil organic matter include cobalt, copper, mercury, nickel and lead (Adriano 2001). A number of studies have systematically investigated soil properties to determine those that exert the greatest controls on the bioaccessibility of arsenic (Cave et al. 2007; Tang et al. 2006; Yang et al. 2005). Although these studies cannot claim to have covered all soil types and soil properties, none of the outcomes showed that organic carbon was significantly influencing the bioaccessibility of arsenic.

Besides the direct complexation of contaminants by organic carbon there are also two other important secondary factors which affect contaminant bioaccessibility:

- redox conditions;
- organic matter competition for sorption of contaminants on oxides and clay.

Several studies (e.g. Baker et al. 2003; Chen et al. 2003; Lindsay 1991; Rose et al. 1990) show that organic matter has an abiotic and biotic role in the reduction of Fe-oxides, causing dissolution of the host oxide and release of adsorbed contaminants. In addition, organic matter can compete with contaminants for adsorption sites causing displacement from the oxide matrix into more available forms (Dixit and Hering 2003). An example of this (Wragg 2005) is shown in soils containing arsenic derived from natural underlying Jurassic ironstone. A small but statistically significant increase in the bioaccessibility of arsenic in garden soils compared to rural soils was found, which was attributed to gardening practices including the addition of organic matter to improve soil fertility. Stewart et al. (2003b) showed that the bioaccessibility of chromium(VI) was significantly influenced by reduction processes catalyzed by soil organic carbon. Other studies show how organic matter mediates the adsorption of contaminant to clay minerals (Cornu et al. 1999; Lin and Puls 2000; Luo et al. 2006; Manning and Goldberg 1996). In addition, although further discussion is beyond the scope of this chapter, soil organic matter is known to sequester organic contaminants therefore playing a key role in the potential reduction of the bioaccessibility of any organic compounds present (Ruby et al. 2002).

7.2.3 Mineral Constituents

The inorganic constituents of soils usually make up the majority of the mass of the soil and it is the interaction of contaminants with the surfaces of these materials that is a major control on bioaccessibility. Davis et al. (1996b) studied the mineralogical constraints on the bioaccessibility of arsenic in mining sites. They concluded that the arsenic bioaccessibility compared to the total arsenic content in the soils was constrained by:

- encapsulation in insoluble matrices, e.g. energite in quartz;
- formation of insoluble alteration or precipitation rinds, e.g. authigenic iron hydroxide and silicate rinds precipitating on arsenic phosphate grains; and
- formation of iron oxide and arsenic oxide and arsenic phosphate cements that reduce the arsenic–bearing surface area available for dissolution.

In a previous study on lead in Montana soils, Davis et al. (1993) found similar results in which the solubility was constrained by alteration and encapsulation which limited the available lead-bearing surface area. Ruby et al. (1996) diagrammatically summarised how the chemical and mineralogical forms of arsenic and lead relate to their bioaccessibility. Figure 7.7 shows the possible physico-chemical processes governing the bioaccessibility of lead at a contaminated site.

Whilst these early studies provided a good insight into the factors governing arsenic and lead bioaccessibility they were very much aimed at soils from mining areas where the contaminants were introduced into the soils as products from ore processing.

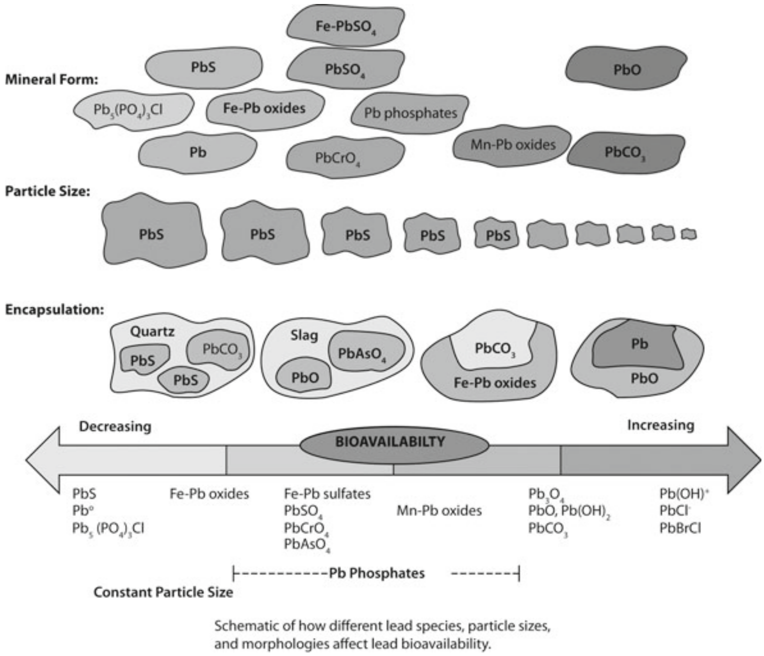


Fig. 7.7 Schematic diagram of how different lead species, particle size and morphologies affect lead bioavailability (after Ruby et al. 1996)

In the UK, a number of studies have examined the bioaccessibility of elevated concentrations of arsenic in soils developed over Jurassic ironstones (Cave et al. 2003, 2007; Palumbo-Roe et al. 2005; Wragg 2005; Wragg et al. 2007). In this instance, there was a highly significant correlation between the total arsenic content and the total iron content, but there was no significant correlation between the bioaccessible arsenic and the total iron content. Palumbo-Roe et al. (2005) concluded that in these soils the bioaccessible arsenic is mainly contained within calcium iron carbonate (sideritic) assemblages and only partially in iron aluminosilicates, probably berthierine, and iron oxyhydroxide phases, probably goethite. It is suggested that the bulk of the non-bioaccessible arsenic is bound up with less reactive more highly crystalline (see Fig. 7.8) iron oxide phases.

These studies highlight the role of two very important mineral classes which are ubiquitous in soils and have been shown to be key controls on bioaccessibility. These are:

- clays;
- oxides of iron, manganese and aluminium.

Clay minerals are produced through hydrolysis weathering reactions which is the reaction between hydrogen ions and an aluminosilicate mineral (such as feldspar) to form soluble cations, silicic acid and a clay mineral. They are characterised by two-dimensional sheets of corner sharing SiO_4 and AlO_4 tetrahedra. These tetrahedral sheets have the chemical composition $(AlSi)_3O_4$, and each tetrahedron shares

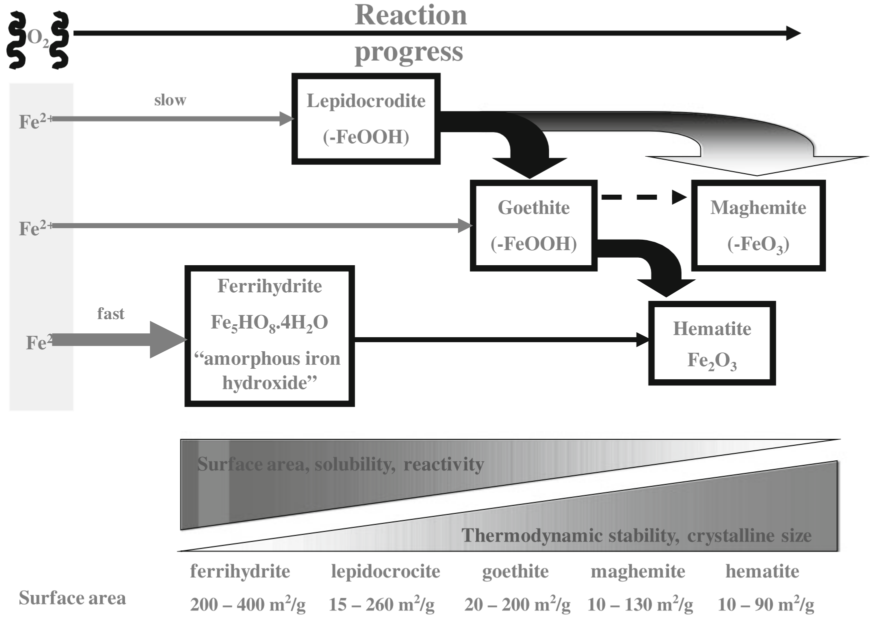


Fig. 7.8 Diagrammatic representation of the ageing processes of Fe oxides in soils

three of its vertex oxygen atoms with other tetrahedra forming a hexagonal array in two-dimensions. The fourth vertex is not shared with another tetrahedron and all of the tetrahedra ‘point’ in the same direction, i.e. all of the unshared vertices are on the same side of the sheet.

Oxides of iron, manganese and aluminium are often referred to as hydrous oxides and like clays are principally derived from weathering reactions of rock minerals. Although different in chemical structure, these two classes of minerals are very fine grained ($<2\ \mu m$) and hence have a very large reactive surface area and similar modes of action in binding contaminants and hence controlling their bioaccessibility. These modes are:

- cation and anion exchange;
- specific adsorption.

For ion exchange, contaminant ions are bound electrostatically to the clay or oxide surface sites with an opposite charge. As already discussed in the organic carbon section, organic matter can also act as ion exchangers. A measure for the ability of the soil to attract and retain cations is known as the cation exchange capacity (CEC). In general, oxides contribute little to the CEC when soil pH is <7, under these conditions, the main contribution comes from organic matter and clays. Anion exchange occurs where negatively charged ions are attracted to positively charged sites. The highest anion exchange for oxides occurs at low pH. Cation exchange is reversible, diffusion controlled and stoichiometric and has an order of selectivity based on the size, concentration and charge of the ion. Electrostatically bound contaminants are displaced relatively easily from the soil matrix in the presence of the low pH conditions of the human stomach.

Specific adsorption involves the exchange of cations and anions with surface ligands on solids to form partly covalent bonds with lattice ions. As with ion exchange, the process is highly dependant on pH, charge and ionic radius. In contrast to ion exchange, however, contaminants bound by this mechanism are far less labile. Brummer (1986) showed that the sorptive capacities of iron oxides and aluminium oxides were up to 26 times higher than their ionic complexes at pH 7.6.

Many studies have confirmed the importance of clays and oxides on the bioaccessibility of contaminants, (e.g. Ahnstrom and Parker 2001; Boonfueng et al. 2005; Bowell 1994; Cave et al. 2007; Chen et al. 2002; Esser et al. 1991; Foster et al. 1998; Garcia-Sanchez et al. 2002; Lin et al. 1998; Manceau et al. 2000; Matera et al. 2003; Palumbo-Roe et al. 2005; Smith et al. 1998; Somez and Pierzynski 2005; Stewart et al. 2003a; Sultan 2007; Violante and Pigna 2002; Violante et al. 2006; Yang et al. 2005; Zagury 2007). It is clear that iron oxides are most commonly reported as hosts for sorbed arsenic (Camm et al. 2004; Cances et al. 2005; Cave et al. 2007; Palumbo-Roe et al. 2005; Wragg et al. 2007). Depending on the form of the iron oxide present in the soil they can be both sources of bioaccessible and non-bioaccessible arsenic. Figure 7.8 shows how progressive ageing of iron oxides from amorphous forms through to more crystalline forms increases the thermodynamic stability and, hence, specifically adsorbed contaminants, notably arsenic, are less easily mobilised. For other metals the picture is less clear and the geochemical host of the contaminant under study is very dependant on the history of the soil.

7.2.4 Solid Phase Speciation and Bioaccessibility

Adsorption of contaminants to different solid phases has been shown to be a key factor in determining the bioaccessibility. It is, therefore, very important that analytical methodologies are available that can be used to measure the physico-chemical forms of contaminants in the soil. Such methods may then provide information on the potential environmental redistribution of contaminants under different soil conditions, and ultimately be used as additional lines of evidence to support in-vitro bioaccessibility testing in the assessment of human health risks from soil ingestion.

Spectroscopic methods such as x-ray absorption fine structure (XAFS) and x-ray absorption near edge structure (XANES), that directly measure the oxidation state and chemical bonds holding the contaminant in the soil, have been used very successfully, (e.g. Cances et al. 2005; Cutler et al. 2001; Manceau et al. 2000; Peak et al. 2006; Welter et al. 1999). However, these methods require the use of a synchrotron source and can be expensive and time consuming. A relatively simple and well-adopted method to assess metal pools of differential relative lability in soils is the use of sequential extraction with reagents of increasing dissolution strength. Each reagent should target a specific solid phase associated with the contaminant. Many of these extraction schemes have been described in the literature and have been reviewed in Filgueiras et al. (2002). In many instances, the steps with low dissolution strength are equated to the bioaccessible fraction or are used along side dedicated bioaccessibility tests to help interpret the geochemical source of the bioaccessible fraction (Datta et al. 2006; Denys et al. 2007; Jimoh et al. 2005; Liu and Zhao 2007a, b; Marschner et al. 2006; Palumbo-Roe et al. 2005; Reeder et al. 2006; Schaider et al. 2007; Siebielec et al. 2006; Tang et al. 2004, 2006, 2007).

Cave et al. (2004), amongst other workers, has highlighted major shortcomings with traditional sequential extraction methods and developed a new procedure called Chemometric Identification of Substrates and Element Distributions (CISED). This procedure uses increasing strengths of simple mineral acids as the extractant, followed by chemometric data processing of the resulting multi-element data obtained from the extract analysis. This method has been shown to work well for a number of contaminants in the NIST 2710 reference soil compared to more traditional sequential extraction schemes and has subsequently been applied very successfully in a number of studies to identify the source of bioaccessible contaminants in soils (Cave et al. 2003; Palumbo-Roe and Klinck 2007; Palumbo-Roe et al. 2005; Wragg 2005; Wragg et al. 2007).

In addition to the solid phase distribution of contaminants, their chemical form (speciation) can affect bioaccessibility and also the toxicity. In human Exposure Assessment, if a bioaccessibility factor is applied to the total soil concentration, the soil speciation needs to be understood. For redox sensitive contaminants such as arsenic, antimony or chromium, which have different toxicity levels according to their redox state, this verification is particularly important. Denys et al. (2009) investigated whether the speciation of antimony changed during the human gastro-intestinal digestion process in four soils, sampled from a former lead-mining site using the differential pulse anodic stripping voltammetry DPASV technique. The results showed that for each soil sample, as the pentavalent form was present and no change in speciation occurred during the digestion process, a resulting Risk Assessment would result in no additional human health risks due to changes in speciation.

7.2.5 Soil Ageing

As well as short term fluctuations, soil can undergo longer term changes caused by changes in land use or other environmental factors such as acid rain, flooding and global warming. As such it is important to note that when the bioaccessibility of a contaminant within a soil is assessed it only applies to a snapshot in time and that bioaccessibility can change with time. A number of studies have looked at how the bioaccessibility of freshly contaminated soils changed with ageing over relatively short timescales (periods of months). For arsenic, the general view is that the bioaccessibility of freshly contaminated soil decreases with time. This is thought to be due to oxidation of more soluble arsenic(III) forms to less soluble arsenic(V), followed by sorption onto Fe-oxides and Fe-oxyhydroxides (Datta et al. 2007; Fendorf et al. 2004; Juhasz et al. 2008; Lin and Puls 2000; Lombi et al. 1999; Tang et al. 2007; Yang et al. 2002, 2003, 2005). The absolute magnitude of the effect, however, varied significantly between soil types. Other studies on cadmium and lead (Lee 2006; Tang et al. 2006; Yang et al. 2003) suggest that the decrease in bioaccessibility for these metals is not as marked as for arsenic, although this was very specific to soil type and pH. Stewart et al. (2003b) showed chromium bioaccessibility decreased with duration of exposure, with aging effects being more pronounced for chromium(III). The decrease in chromium bioaccessibility was rapid for the first 50 days and then slowed dramatically between 50 and 200 days. In general, the effects of chromium solid phase concentration on bioaccessibility was small, with chromium(III) showing the most pronounced effect, higher solid phase concentrations resulted in a decrease in bioaccessibility. Chemical extraction methods and X-ray Adsorption Spectroscopy analyses suggested that the bioaccessibility of chromium(VI) was significantly influenced by reduction processes catalyzed by soil organic carbon.

7.2.6 Statistical Modelling of Bioaccessibility

Yang et al. (2002) carried out a detailed study of arsenic contaminated soils, in which the arsenic originated from processes other than mining. Their studies showed that soils with lower soil pH and higher Fe-oxide content exhibited lower bioaccessibility and were able to model the bioaccessible arsenic content using these factors. However, the model was not able to accurately predict the bioaccessibility of arsenic in a different set of contaminated soils, previously used in an independent

Cebus monkey dosing trial, consistently overpredicting the bioavailability, resulting in an unacceptably large uncertainty. Juhasz et al. (2007) found that they could model the in-vitro bioaccessibility of soils contaminated by arsenic from herbicides, pesticides and mining waste, using the total arsenic content and total or dithionitecitrate extractable (free) iron. However, the bioaccessible content of a naturally mineralised site could not be modelled in this way. In a quite different approach a self-modelling mixture algorithm was used (Cave et al. 2007) to convert raw Near Infra-Red (NIR) spectra of soils, developed over Jurassic ironstones, into five underlying spectral components and associated coefficients. The five spectral components were shown to be significantly correlated to the total arsenic, bioaccessible arsenic and total Fe-contents of the soils and tentatively assigned to crystalline Fe oxides, Fe oxyhydroxides and clay components in the soils. A linear regression model, using the spectral component coefficients associated with the clay fraction, the Fe oxyhydroxides and the total arsenic content of the soils as independent variables, was shown to predict the bioaccessible arsenic content of the soils, as measured by an in-vitro laboratory test, with a 95% confidence limit of ± 1.8 mg kg^{-1} and a median R^2 value of 0.80.

7.2.7 Soil Sampling and Preparation for Bioaccessibility/Bioavailability Measurements

In order for the final results of bioaccessibility/bioavailability testing to be meaningful the samples under investigation need to be representative of the sampling location and the grain size applicable to the resulting Human Health Risk Assessment. As such, a number of sampling and preparation considerations need to be met as part of the underlying analytical protocol.

Prior to soil sampling, consideration should be given to:

- the history of the location – including the effect of the local geology and or historical anthropogenic contamination and the potential effect of contaminant 'hotspots';
- the equipment to be used to collect the samples – clean, high quality sampling tools and containers in order to avoid sample contamination;
- the number of samples to be collected per averaging zone; Nathanail (2009) recommend a minimum of 10 samples in order to 'gain and adequate appreciation of the variation in bioaccessibility'. However, when considering in-vivo bioavailability measurements, the cost of testing may be the driving force in determining the number of samples, per location, to be collected;
- the depth at which samples are collected – surface soil samples from 0 to 15 cm depth, as these are representative of the material to which humans are likely to be exposed (Cave et al. 2003);
- the variation of soil composition and resulting bioaccessibility across site, therefore separate bioaccessibility sampling areas will be required for each soil type represented;

- the type of sample to be collected – grab or composite? Composite sampling is the standard practice for geochemical surveying work (Kelleher 1999). Composite sampling is analogous of making three replicate measurements of analytical data and averaging the data point (Wragg 2005). However, 'hotspots' may be missed or the contaminant concentration may be reduced as not all samples to be composited have the same/similar contaminant concentration;
- the preparation of the individual samples – samples should be dried at <35°C and gently disaggregated, but never crushed, in order to break up large clasts and homogenized. A representative portion of the bulk material should be sub-sampled and sieved to <250 μm, as this fraction is considered to be the upper limit of particle sizes that are likely to adhere to children's hands (often the at risk receptor) (Duggan et al. 1985), and tested for its total and bioaccessible/bioavailable contaminant content.

7.3 Considerations for the Potential Use of Site Specific Bioaccessibility Measurements

Bioaccessibility measurements are not necessarily applicable to all contaminants, all soil types and all Risk Assessment scenarios. The following key questions should be considered before embarking on bioaccessibility testing:

- Is the contaminant concentration slightly above the guidance value for the soil under consideration?
    - Research in the UK (Cave et al. 2003; Nathanail and Smith 2007; Palumbo-Roe and Klinck 2007; Palumbo-Roe et al. 2005; Wragg et al. 2007) has shown that specific soil types have a well defined distribution of % bioaccessibility values. Using this knowledge it is then possible to estimate the maximum total soil concentration, for a soil type, where bioaccessibility data would assist the assessment of risk. For example, in the Jurassic Ironstone soils in Lincolnshire (Eastern England) the modal bioaccessible arsenic fraction is approximately c. 10%. Working on a possible arsenic guideline value of 20 mg kg^{-1}, this would suggest that a total arsenic soil concentration of up to 200 mg kg^{-1} would be suitable for using bioaccessibility testing in a further detailed quantitative Risk Assessment. Whereas, dependant on the local federal guidelines (i.e. Superfund site etc) a soil with a concentration of 1,000 mg kg^{-1} arsenic is unlikely to be suitable for bioaccessibility testing. If the concentration is very high, this will override all subsequent points and bioaccessibility will no longer be an option.
- Is remediation likely to be very expensive, unsustainable or not technically feasible?
    - For point source contamination, it is likely that remediation will only involve a relatively small volume of contaminated material. However, for diffuse contamination, particularly from natural geogenic sources, e.g. naturally

mineralised soils in Devon and Cornwall in the south-west of England, it is practically impossible to remediate in any sensible fashion. The lines of evidence approach, which includes bioaccessibility testing and contaminant solid phase distribution determination with associated data interpretation, may be a pragmatic way forward (Palumbo-Roe and Klinck 2007).

- Is there an adverse environmental risk associated with remediation?
  - Remediation can involve a large amount of heavy machinery and transport (plant), which will have a significant effect on the quality of life to the surrounding population and adverse health effects through dust inhalation;
- Is the number of contaminants driving the risk three or fewer?
  - If the site is contaminated with a wide variety of contaminants, consideration must be given to which contaminant is driving the risk calculation. As a rule of thumb, if three or less contaminants are causing a potential risk then bioaccessibility will probably decrease the estimates of risk.
- Is there in-vivo validation data associated with the in-vitro method under consideration?
  - The criteria associated with validation of a bioaccessibility test are discussed in detail under Section 7.1.3.
- Does the local regulator accept in-vitro bioaccessibility in Human Health Risk Assessment?
  - The regulators responsible for a given site should be contacted prior to a site investigation in order to determine whether bioaccessibility testing is accepted. This will ensure that financial resources are not wasted by unnecessary testing, in the case of that bioaccessibility data is not acceptable, and natural resources such as soil is not unnecessarily remediated, if the opposite is true.
- Is the land use likely to change in the future?
  - The application of bioaccessibility data needs to be considered in terms of the land use at the time of the assessment and any proposed future land use. Because of soil ageing and weathering etc. (see Section 7.2.5) it is considered that naturally occurring contaminants in soil will have a lower bioaccessibility/bioavailability than those found in made ground. In addition the bioaccessibility of contaminants may be altered (increased or decreased) by land practices such as the liming of soil to raise the pH or the addition of organic matter (Wragg 2005).

7.4 Examples of Bioaccessibility Studies

Bioaccessibility research is a rapidly developing scientific area, as such there are many examples of its application in the scientific literature. It is beyond the scope of this chapter to give a comprehensive review, however the following section gives examples of bioaccessibility related studies for a variety of contaminants arising from geogenic sources and anthropogenic influences. Other examples of where

Table 7.1 Examples of where bioaccessibility testing has been applied to contaminated soils (2003–2009)

Authors	Year	Source/influence	Contaminants
Siebielec et al.	2006	Geogenic	Lead and Cadmium
Cave et al.	2007	Geogenic	Arsenic
Juhasz	2007	Geogenic and Anthropogenic	Arsenic
Beak et al.	2008	Geogenic	Lead
Madrid et al.	2008	Geogenic	Copper, Nickel, Lead and Zinc
Smith et al.	2008	Geogenic	Arsenic
Yang et al.	2003	Anthropogenic	Lead and Arsenic
Sonmez and Pierszynski	2005	Anthropogenic	Lead
Yang et al.	2005	Anthropogenic	Arsenic
Nico et al.	2006	Anthropogenic	Arsenic and Chromium
Schwab et al.	2006	Anthropogenic	Lead
Pouschat and Zagury	2006	Anthropogenic	Arsenic
Tang et al.	2006	Anthropogenic	Cadmium
Gal et al	2007	Anthropogenic	Arsenic and Antimony
Gron et al	2007	Anthropogenic	Polyaromatic hydrocarbons
Ljung et al	2007	Anthropogenic	Arsenic, Cadmium, Chromium. Nickel and Lead
Subacz et al.	2007	Anthropogenic	Arsenic
Van de Wiele et al.	2007	Anthropogenic	Lead
Roberts et al.	2007	Anthropogenic	Arsenic
Bruce et al.	2007	Anthropogenic	Arsenic and Lead
Sarkar et al.	2007a, b	Anthropogenic	Arsenic
Schaider et al.	2007	Anthropogenic	Zn, Lead and Cadmium
Morrison and Gulson	2007	Anthropogenic	Arsenic, Cadmium, Lead
Finzgar et al.	2007	Anthropogenic	Lead and Zinc
Mosely et al.	2008	Anthropogenic	Lead
Romero et al.	2008	Anthropogenic	Lead
Bosso et al.	2008	Anthropogenic	Lead
Bosso and Enzwiler	2008	Anthropogenic	Lead
Rasmussen et al.	2008	Anthropogenic	Copper and Zinc
Poggio et al.	2009	Anthropogenic	Lead, Copper, Zinc, Nickel and Chromium
Stewart et al.	2003a, b	Spiked	Chromium
Juhasz et al.	2008	Spiked	Arsenic

bioaccessibility testing has been applied to contaminated soils over the period 2003 to 2009 have been listed in Table 7.1.

7.4.1 Geogenic Sources

Palumbo-Roe et al. (2005) and Wragg et al. (2007) have studied the bioaccessibility in soils with elevated arsenic concentrations (up to circa 200 mg kg^{-1}). A combination of a physiologically based extraction test combined with geochemical testing (using either the CISED methodology described in Section 7.2.4 or multivariate statistical modelling of geochemical soil survey data) was used. The findings showed that the bioaccessible arsenic was generally less than and mainly contained within calcium iron carbonate (sideritic) assemblages and only partially iron aluminosilicates, probably berthierine, and iron oxyhydroxide phases, probably goethite. The bulk of the non-bioaccessible arsenic was bound up with less reactive iron oxide phases.

Naturally occurring arsenic in soils (13–384 mg kg^{-1}) at a new housing site in southwest England (Nathanail and Smith 2007) were demonstrated not to pose unacceptable risk to human health by site specific estimates of bioavailability and region specific estimates of soil to plant uptake factors. Independent lines of evidence, consisting of data from sequential extraction of representative test soils and soil to plant uptake factors for the site were used to justify the arsenic exposure factors for oral bioavailability. The results of the study and subsequent Risk Assessment avoided the need for remediation and unnecessary public concern.

Denys et al. (2007) studied the bioaccessibility of lead in soils with a high lead carbonate (cerrusite) content (up to 870 g kg^{-1}). The authors found that lead bioaccessibility, using the RIVM in-vitro protocol, in high carbonate soils can be low (down to 20% of the total soil Pb content) and is not correlated with cerussite soil contents even if the concentration of this mineral is relatively high. This research indicates that mineralogical analysis alone is not a reliable predictor of the bioaccessible fraction.

7.4.2 Anthropogenic Influences

Basta and co-workers have investigated the bioaccessibility of arsenic and lead, using the in-vitro gastro intestinal method (IVG) in smelter impacted soils over a number of years (Basta et al. 2007; Rodriguez et al. 1999). Method development of the IVG included validation and dosing trials of immature swine for fifteen soils ranging in total arsenic concentrations of 401–17,460 mg As kg^{-1}. The research indicates that IVG methodology was linearly correlated with the animal model ($r = 0.83$, $p < 0.01$) for arsenic relative bioavailability ranging from 2.7 to 42.8% (Rodriguez et al. 1999). Further research by this group investigated the effect of the dosing vehicle used to determine the bioaccessibility, using the same contaminated materials as employed in the 1999 research (Basta et al. 2007). Similarly to previous research, the IVG was shown to produce strong relationships with the in-vivo arsenic bioavailability, with or without the presence of a dosing vehicle ($r = 0.92$ (with), $p < 0.01$) and $r = 0.96$ (without) $p < 0.01$). Further to the work on arsenic, the applicability of the IVG methodology has been investigated for soils contaminated with cadmium (Schroder et al. 2003). Relative bioavailable cadmium for ten soils (containing total cadmium ranging between 23.8 and 465 mg kg^{-1}) was obtained from dosing trials using juvenile swine and ranged from 10.4 to 116%. Linear regression with cadmium bioaccessibility from the IVG methodology indicated strong correlations with both the stomach ($r = 0.86$) and intestinal ($r = 0.80$) phases of the IVG.

In a study of and abandoned copper-arsenic mine in Devon, in the South-west of England in the UK (Devon Great Consols (DGC)) Palumbo-Roe and Klinck (2007) investigated the mineralogical factors that control the bioaccessibility of arsenic in soils influenced by the past mining operation in the area. Bioaccessibility (determined using a physiologically based extraction test, as described by Cave et al. (2003)) was related to the solid phase distribution of arsenic present in the test soils (determined by the CISED method (Cave et al. 2004)). The results of this study show that the mine soils from DGC have higher arsenic bioaccessibility (median 15%) than those not affected by mining activities and other background soils collected from the sampling area (the Tamar catchment), with a median bioaccessibility of 9%. Determination of the solid phase distribution of the arsenic present indicated that the arsenic present in the test soils was mainly hosted by an iron oxyhydroxide component, whose partial dissolution was responsible for the bioaccessible arsenic fraction. The degree of crystallinity of this component was thought to be an important control on the arsenic bioaccessibility.

Caboche et al. (2008) investigated the bioaccessibility at sites influenced by different sources of historical contamination: a former lead-zinc mining site and a site in which soils were contaminated by atmospheric deposits of lead containing particles. The total concentrations ranged from 48 to 247 $\mathrm{mg{\cdot}kg_{dw}}^{-1}$ for arsenic and from 1,462 to 16,267 $\mathrm{mg{\cdot}kg_{dw}}^{-1}$ for lead, with the most contaminated soils representing the mining site. Bioaccessibility data for the two contaminant sources showed that for the mining soils arsenic and lead bioaccessibility was lower (ranging from 2.7 to 8% for arsenic and 13.2 to 35.5% for lead) compared to 10 to 52% for the soils contaminated by atmospheric deposition. The results of this study underline the importance of bioaccessibility data when comparing the exposure to contaminants present at contaminated sites.

7.5 The BARGE Network

The Bioaccessibility Research Group of Europe (BARGE) is a network of European institutes and research groups, formed in 1999/2000, to ‘study human bioaccessibility of priority contaminants in soils’ (www.bgs.ac.uk/barge). A priority objective of the network is the provision of oral bioaccessibility/bioavailability data for Human Health Risk Assessments and policy making that is both robust and defensible.

A key driver for the inception of the network, identified by representatives of the 16 Member States of the concerted action CLARINET (Contaminated Land Rehabilitation Network for Environmental Technologies), was a European wide urgency for more realistic oral bioavailability factors that could be used in site specific Risk Assessment and policy making. Initial collaborative efforts, financially supported by the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM) and the participating institutes, allowed for the bringing together of a multi-disciplinary team of research scientists (with a diverse range knowledge base in pharmaco-kinetics, physiology, geochemistry and analytical measurement techniques), policy makers and Risk Assessment practitioners, all key in the decision making process.

Since its inception, the BARGE group has been active in the field of oral bioavailability and bioaccessibility. The activities include the undertaking of a number of round robin studies (see section below); liaising with the International Standards Organisation on the preparation of a technical specification entitled 'Soil quality – Assessment of human exposure from ingestion of soil and soil material – Guidance on the application and selection of physiologically based extraction methods for the estimation of the human bioaccessibility/bioavailability of metals in soil'; the dissemination of findings through the development of a website (www.bgs.ac.uk/barge), the hosting of workshops/dedicated speaker sessions on bioaccessibility related topics at International conferences such as ConSoil (http://consoil.ufz.de/); and the preparation of a peer reviewed special publication on bioaccessibility topics (Gron and Wragg 2007).

As the research on oral bioavailability, and the development of oral bioavailability factors, is not an issue specific to Europe, and affects contaminated site practitioners worldwide, the BARGE group is an ever expanding International research network, combining the cross-continental collaborative efforts of Europe and North America. In a similar vein to the BARGE group, researchers, Risk Assessment practitioners and regulators in Canada with an interest in oral bioavailability and the application of resulting tools have joined forces to form BioAccessibility Research Canada (BARC). The long term aim of this group is 'to provide a scientific basis for evaluating and predicting inorganic and organic contaminant bioaccessibility in soils found at contaminated sites in Canada' (BARC 2006). To address issues relating specifically to the measurement of arsenic oral bioavailability and bioaccessibility by in-vivo and in-vitro methodologies respectively, researchers from both the BARGE and BARC groups have joined forces with other research institutes and government bodies to share knowledge via an electronic forum.

7.5.1 Inter-Laboratory Studies

To date, four international inter-laboratory studies have been carried out to investigate different aspects of bioaccessibility in the human gastro-intestinal tract. Three have been undertaken by the BARGE group, two of which have been published in peer reviewed publications (Oomen et al. 2002; Van de Wiele et al. 2007). Oomen et al. (2002) described a multi-laboratory study which compared the bioaccessibility data returned by five in-vitro methods for three solid materials, contaminated with arsenic, cadmium and lead. The methods included physiologically and non-physiologically based test systems of dynamic and static nature. The salient points of the study were that, in many cases the bioaccessibility was <50%, an important factor for Human Health Risk Assessment, although a wide range of bioaccessibility values were observed across the in-vitro methods studied. The primary driver for the range in bioaccessible values was attributed to the difference in gastric pH of the various test systems. High bioaccessibilities were typically observed for the simplest method (stomach compartment only) with a low gastric pH and the lowest values for the system with the highest gastric pH (4.0). Low bioaccessible values, however, were also observed with two-phase systems incorporating low gastric pH and an intestinal compartment at neutral pH, indicating that pH plays an important role in both phases of the in-vitro test.

The multi-laboratory comparison study by Van de Wiele et al. (2007) assessed the bioaccessibility of soil-bound lead under both fed and fasted conditions in a number of in-vitro test systems. The study utilized both the soils and the resulting lead bioavailability data (Maddaloni et al. 1998), as a reference point for comparison, from a previous human in-vivo trial. As with the previous study by the BARGE group, both static and dynamic in-vitro methodologies were included, but unlike the previous study, all test systems were physiologically-based. The study showed that regardless of the nutritional status of the model (fed or fasted state), the lead bioaccessibility was significantly different ($p < 0.05$) between the methods with bioaccessible lead ranging between ca. 2 to 35%. Comparison with the available in-vivo data indicated that the simulation of the human gastrointestinal system under fed conditions overestimated lead bioavailability, whilst under fasted conditions a number of the models investigated underestimated oral bioavailability (Van de Wiele et al. 2007). This study and the previous study by Oomen et al. (2002) note that differences in the measured bioaccessibilities are often due to key methodological parameters such as gastric and intestinal pH. This further study notes that the method of separation (centrifugation, filtration, ultrafiltration) used in the test system is a critical factor in separating the bioaccessible from the non-bioaccessible fraction present and in the interpretation of the results for use in Human Health Risk Assessment.

The third inter-laboratory study initiated by the BARGE group had the goal of progressing development of a harmonised bioaccessibility methodology in order to carry in-vitro testing for Human Health Risk Assessment of contaminated soils. The collective efforts of the network modified the previously published RIVM physiologically based in-vitro method and the trial determined the analytical performance characteristics of the bioaccessibility measurement (repeatability and reproducibility). Modifications were made in order that the methodology was robust and provided adequate conservatism, at least for first tier Risk Assessment, for future use across the local geological conditions of the individual member countries (Wragg et al. 2009). The study utilized slag materials, river sediments, soil material and house dusts that had been previously investigated for their arsenic, cadmium and lead bioaccessibility and oral bioavailability contents by Professors Nick Basta (Basta et al. 2007; Rodriguez et al. 1999; Schroder et al. 2003, 2004a) and/or Professor Stan Casteel. The first inter-laboratory study showed that the harmonised methodology had the potential to meet the benchmark criteria set by BARGE. In addition, with the aid of further testing on soils that are representative of the local geology to the participating countries and not just highly elevated mining and slag waste, the method could be standardised for international use.

Health Canada has funded the BARC group to assess the variability in the reported data across Canadian laboratories in the form of a simple inter-laboratory comparison. This initial BARC study includes domestic commercial and research facilities and international research laboratories from the UK, the Netherlands and the US. Future round robin studies are planned by the group in order to examine the variability amongst individual methods and by comparison of results to toxicological reference values (BARC 2008).

7.5.2 Utilization of Bioaccessibility Data Across Europe

The utilization of the concept of oral bioavailability/bioaccessibility is used in a highly variable manner across Europe. In France, until recently, bioaccessibility was not taken into account in Human Health Risk Assessment. However, more recently, the Institut National de l'Environnement Industriel et des Risques (INERIS) has been tasked with the development of research programs that include the application of bioaccessibility testing in Human Health Risk Assessment and assessing the benefits of the inclusion of such testing regimes. As a part of its risk based approach to the management of contaminated land there is guidance is in place in the UK that recognises the potential benefits of, and allows for the inclusion of, oral bioavailability/bioaccessibility data in Human Health Risk Assessments when addressing exposure to significant possibilities of significant harm (SPOSH) (Department for the Environment Food and Rural Affairs and the Environment Agency 2002; ODPM 2004). Similarly, in Canada, although absorption factors for ingestion are usually 100% in screening level Risk Assessments, oral bioavailability is often determined as bioaccessibility, and for complex Risk Assessments site specific bioaccessibility values may be determined as a surrogate of bioavailability (Health Canada 2004). However, currently there is a paucity of guidance available concerning the application of bioaccessibility data in UK and Canadian Risk Assessments, which has lead to a degree of uncertainty within the practitioner community regarding the use of this potentially beneficial tool. In the US, the 'Risk Assessment Guidance for Superfund, Volume 1' (USEPA 1989) considers the determination of bioavailability on a site-specific basis, and more recently guidance has become available for a number of issues surrounding site-specific bioavailability data (USEPA 2007a). Such issues include:

- implementation of a validated methodology, and documentation of data collection and analysis;
- EPA criteria for the evaluation of a bioavailability methodology, to test its suitability for use.

Additionally, after recent validation of an in-vitro methodology for lead against in-vivo swine data, there is growing support for the use of reliable bioaccessibility data, in the first instance for Pb in Human Health Risk Assessments (USEPA 2007b).

Despite apparent confusion surrounding the science and application of bioaccessibility testing within the UK contaminated site community, process understanding of the mechanisms controlling bioaccessibility has continued with respect to priority UK contaminants (arsenic, chromium, nickel, lead and poly aromatic hydrocarbons (PAHs)), in addition to the application of bioaccessibility data in site-specific UK detailed quantitative Risk Assessments. Research on bioaccessibility testing has focussed on assessing the solid phase distribution of contaminants, such as arsenic, in soils to explain the bioaccessible sources of this contaminants (Cave et al. 2003; 2007; Nathanail and Smith 2007; Nathanail et al. 2004, 2005a, b; Palumbo-Roe and Klinck 2007; Palumbo-Roe et al. 2005). Inclusion of the physico-chemical sources of the bioaccessible contaminant can provide a better understanding of the mechanisms by which the potentially harmful contaminant becomes bioaccessible and may aid the process understanding of how the bioaccessibility may change with time and changes in land use. Nathanail et al. (2005a) included in-vitro bioaccessibility data, as one line of evidence, to a) assist in the derivation of site specific assessment criteria and b) demonstrate that the pedogenic arsenic present at a dwelling in Wellingborough in the UK, did not pose an unacceptable human health risk to the owners. Currently there is no official status regarding bioaccessibility testing, in the Netherlands. In 2006 the RIVM provided the Dutch government with a report recommending the use of bioaccessibility testing for lead on site specific basis. In the latest revision of the Dutch Soil Quality Standards (Intervention Values) (Ministry of VROM 2008) a relative bioavailability factor for lead of 0.73 has been applied). The Danish Environmental Protection Agency (EPA) has allowed for the use of oral bioavailability data for lead and cadmium (in the form of in-vitro bioaccessibility data) in the evaluation of sites in compliance with Soil Quality Standards. As part of this allowance, the Danish EPA has stipulated that oral bioavailability data must be obtained using the RIVM fasted state methodology. For lead the data must be generated using the stomach phase and for cadmium the intestinal phase should be included. In Flanders, bioaccessibility testing is considered for the assessment of risk to human health risk from soil contaminated with PAHs and other contaminants (Van de Wiele T 2008; Wragg, J. Nottingham ‘personal communication’).

References

Abrahams PW, Follansbee MH, Hunt A, Smith B, Wragg J (2006) Iron nutrition and possible lead toxicity: an appraisal of geophagy undertaken by pregnant women of UK Asian communities. Appl Geochem 21:98–108

Adriano DC (2001) Trace elements in terrestrial environments, 2nd edn. Springer, New York

Ahnstrom ZAS, Parker DR (2001) Cadmium reactivity in metal-contaminated soils using a coupled stable isotope dilution-sequential extraction procedure. Environ Sci Technol 35:121–126

Amshead HDeW, Graff DJ, Ashmead HH (1985) Intestinal absorption of metal ions and chelates. CC Thomas (ed) Springfield, USA

Baker S, Herrchen M, Hund-Rinke K, Klein W, Kordel W, Peijnenburg W, Rensing C (2003) Underlying issues including approaches and information needs in risk assessment. Ecotoxicol Environ Safety 56:6–19

BARC (2006) BARC newsletter 1. Wragg J, Reimer K, Kingston, ON, BARC

BARC (2008) BARC newsletter winter (2008) Reimer K, Koch I, Paquin V, Wragg J, Kingston, ON, BARC

Barth E, Sass B, Chattopadhyay S (2007) Evaluation of blast furnace slag as a means of reducing metal availability in a contaminated sediment for beneficial use purposes. Soil Sediment Contamin 16:281–300

Basta NT, Foster JN, Dayton EA, Rodriguez RR, Casteel SW (2007) The effect of dosing vehicle on arsenic bioaccessibility in smelter-contaminated soils. J Environ Sci Health Part A 42: 1275–1281

Beak DG, Basta NT, Scheckel KG, Traina SJ (2006) Bioaccessibility of arsenic bound to corundum using a simulated gastrointestinal system. Environ Chem 3:208–214

Beak DG, Basta NT, Scheckel KG, Traina SJ (2008) Linking solid phase speciation of Pb sequestered to birnessite to oral Pb bioaccessibility: implications for soil remediation. Environ Sci Technol 42:779–785

Boonfueng T, Axe L, Xu Y (2005) Properties and structure of manganese oxide-coated clay. J Colloid Interface Sci 281:80–92

Bosso ST, Enzwiler J (2008) Bioaccessible lead in soils, slag and mine wastes from an abandoned mining district in Brazil. Environ Geochem Health 30:219–229

Bosso ST, Enzweiler J, Angelica RS (2008) Lead bioaccessibility in soils and mine waste after immoblization with phosphate. Water, Air Soil Pollut 195:257–273

Bowell RJ (1994) Sorption of arsenic by iron-oxides and oxyhydroxides in soils. Appl Geochem 9:279–286

Bridges CC, Zalups RK (2005) Molecular and ionic mimicry and the transport of toxic metals. Toxicol Appl Pharmacol 204:274–308

Bruce S, Noller B, Matanitobua V, Ng J (2007) in vitro physiologically based extraction test (PBET) and bioaccessibility of arsenic and lead from various mine waste materials. J Toxicol Environ Health-Part A-Curr Issues 70:1700–1711

Brummer GW (1986) The importance of chemical speciation in environmental processes Springer, Berlin, pp 169–192

Cabanero AI, Madrid Y, Camara C (2007) Mercury-selenium species ratio in representative fish samples and their bioaccessibility by an in vitro digestion method. Biol Trace Element Res 119:195–211

Caboche J, Denys S, Tack K, Feidt C (2008) Bioaccessibility and speciation of As and Pb in soils contaminated by two distinct anthropogenic activities. Consoil 2008, Milan, Italy

Camm GS, Glass HJ, Bryce DW, Butcher AR (2004) Characterisation of a mining-related arsenic-contaminated site, Cornwall, UK. J Geochem Exploration 82:1–15

Cances B, Juillot F, Morin G, Laperche V, Alvarez L, Proux O, Hazemann JL, Brown GE, Calas G (2005) XAS evidence of As(V) association with iron oxyhydroxides in a contaminated soil at a former arsenical pesticide processing plant. Environ Sci Technol 39: 9398–9405

Casteel SW, Cowart RP, Weis C, Henningsen GM, Hoffman E, Brattin WJ, Starost MF, Payne JT, Stockham SL, Becker SV, Turk JR (1996) A swine model for determining the bioavailability of lead from contaminated media. In: Tumbleon ME, Schook LB (eds) Advances in swine in biomedical research, vol 2. Plenum, New York, 637–646

Casteel SW, Weis CP, Henningsen GM, Brattin WJ (2006) Estimation of relative bioavailability of lead in soil and soil-like materials using young swine. Environ Health Perspect 114: 1162–1171

Cave M, Taylor H, Wragg J (2007) Estimation of the bioaccessible arsenic fraction in soils using near infrared spectroscopy. J Environ Sci Health, Part A 42:1293–1301

Cave MR, Milodowski AE, Friel EN (2004) Evaluation of a method for identification of host physico-chemical phases for trace metals and measurement of their solid-phase partitioning in soil samples by nitric acid extraction and chemometric mixture resolution. Geochem: Exploration, Environ Anal 4:71–86

Cave MR, Wragg J, Palumbo B, Klinck BA (2003) Measurement of the bioaccessibility of arsenic in UK soils. Environ Agency P5–062/TR1

Cavret S, Feidt C (2005) Intestinal metabolism of PAH: in vitro demonstration and study of its impacts on PAH transfer through intestinal epithelium. Environ Res 98:22–32

Chen J, Gu B, Royer RA, Burgos WD (2003) The roles of natural organic matter in chemical and microbial reduction of ferric iron. Sci Total Environ 307:167–178

Chen M, Ma LQ, Harris WG (2002) Arsenic concentrations in Florida surface soils: influence of soil type and properties. Soil Sci Soc Am J 66:632–640

Cornu S, Saada A, Breeze D, Gauthier S, Baranger P (1999) The influence of organic complexes on arsenic adsorption onto kaolinites. Comptes Rendus De L Academie Des Sciences Serie Ii Fascicule a- Sciences De La Terre Et Des Planetes 328:649–654

Cullen WR (2008) Is arsenic an aphrodisiac? The sociochemistry of an element. The Royal Society of Chemistry

Cutler JN, Jiang DT, Remple G (2001) Chemical speciation of arsenic in uranium mine tailings by X- ray absorption spectroscopy. Can J Anal Sci Spectrosc 46:130–135

Datta R, Makris KC, Sarkar D (2007) Arsenic fractionation and bioaccessibility in two alkaline texas soils incubated with sodium arsenate. Archiv Environ Contamin Toxicol 52: 475–482

Datta R, Sarkar D, Sharma S (2006) Arsenic biogeochemistry and human health risk assessment in organo-arsenical pesticide-applied acidic and alkaline soils: an incubation study. Sci Total Environ 372:39–48

Davis A, Bloom NS, Hee SSQ (1997) The environmental geochemistry and bioaccessibility of mercury in soils and sediments: a review. Risk Anal 17:557–569

Davis A, Drexler JW, Ruby MV, Nicholson A (1993) Micromineralogy of mine wastes in relation to lead bioavailability, Butte, Montana. Environ Sci Technol 27:1415–1425

Davis A, Ruby MV, Bloom M, Schoof R, Freeman G, Bergstom PD (1996a) Mineralogic constraints on the bioavailability of arsenic in smelter-impacted soils. Environ Sci Technol 30:392–399

Davis A, Ruby MV, Bloom M, Schoof R, Freeman G, Bergstrom PD (1996b) Mineralogic constraints on the bioavailability of arsenic in smelter-impacted soils. Int J Rock Mech Mining Sci 33:285A

Denys S, Caboche J, Tack K, Delalain P (2007) Bioaccessibility of lead in high carbonate soils. J Environ Sci Health Part A-Toxic/Hazardous Subst Environ Eng 42:1331–1339

Denys S, Tack K, Caboche J, Delelain P (2009) Bioaccessibility, solid phase distribution, and speciation of Sb in soils and in digestive fluids. Chemosphere 74:711–716

Department for the Environment Food and Rural Affairs and the Environment Agency (2002) Assessment of risks to human health from land contamination: an overview of the soil guideline values and related research, CLR7

Dixit S, Hering JG (2003) Comparison of arsenic(V) and arsenic(III) sorption onto iron oxide minerals: implications for arsenic mobility. Environ Sci Technol 37:4182–4189

Drexler JW, Brattin WJ (2007) An in vitro procedure for estimation of lead relative bioavailability: with validation. Human Ecol Risk Assess 13:383–401

Duggan MJ, Inskip MJ, Rundle SA, Moorcroft JS (1985) Lead in playground dust and on the hands of schoolchildren. Sci Total Environ 44:65–79

Ellickson KM, Meeker RJ, Gallo MA, Buckley BT, Lioy PJ (2001) Oral bioavailability of lead and arsenic from a NIST standard reference soil material. Archiv Environ Contamin Toxicol 40:128–135

Esser KB, Bockheim JG, Helmke PA (1991) Trace-element distribution in soils formed in the Indiana Dunes, USA. Soil Sci 152:340–350

Fendorf S, La Force MJ, Li GC (2004) Temporal changes in soil partitioning and bioaccessibility of arsenic, chromium, and lead. J Environ Quality 33:2049–2055

Filgueiras AV, Lavilla I, Bendicho C (2002) Chemical sequential extraction for metal partitioning in environmental solid samples. J Environ Monitor 4:823–857

Finzgar N, Tlustos P, Lestan D (2007) Relationship of soil properties to fractionation, bioavailability and mobility of lead and zinc in soil. Plant Soil Environ 53:225–238

Foster AL, Brown GE, Tingle TN, Parks GA (1998) Quantitative arsenic speciation in mine tailings using X-ray absorption spectroscopy. Am Mineral 83:553–568

Fowles J, Weinstein P, Tseng C-H (2005) Chapter 22, Environmental medicine in essentials of medical geology: impacts of the natural environment on public health. Ed Olle Selinus. Elsevier

Freeman GB, Dill JA, Johnson JD, Kurtz PJ, Parham F, Mathews HB (1996) Comparative absorption of lead from contaminated soil and lead salts by weanling fischer 344 rats. Fundamental Appl Toxicol 33:109–119

Freeman GB, Johnson JD, Killinger JM, Liao SC, Davis AO, Ruby MV, Chaney RL, Lovre SC, Bergstrom PD (1993) Bioavailability of arsenic in soil impacted by smelter activities following oral-administration in rabbits. Fundamental Appl Toxicol 21:83–88

Freeman GB, Johnson JD, Killinger JM, Liao SC, Feder PI, Davis AO, Ruby MV, Chaney RL, Lovre SC, Bergstrom PD (1992) Relative bioavailability of lead from mining waste soil in rats. Fundamental Appl Toxicol 19:388–398

Freeman GB, Johnson JD, Liao SC, Feder PI, Davis AO, Ruby MV, Schoof RA, Chaney RL, Bergstrom PD (1994) Absolute bioavailability of lead acetate and mining waste lead in rats. Toxicology 91:151–163

Freeman GB, Schoof RA, Ruby MV, Davis AO, Dill JA, Liao SC, Lapin CA, Bergstrom PD (1995) Bioavailability of arsenic in soil and house dust impacted by smelter activities following oral administration in cynomolgus monkeys. Fundamental Appl Toxicol 28:215–222

Gal J, Hursthouse A, Cuthbert S (2007) Bioavailability of arsenic and antimony in soils from an abandoned mining area, Glendinning (SW Scotland). J Environ Sci Health Part A 42: 1263–1274

Garcia-Sanchez A, Alvarez-Ayuso E, Rodriguez-Martin F (2002) Sorption of As(V) by some oxyhydroxides and clay minerals. Application to its immobilization in two polluted mining soils. Clay Minerals 37:187–194

Gron C (2005) Test for bioaccessibility of metals and PAH from soil. DHI Water and Environment, Denmark, 52339/06

Gron C, Wragg J (2007) Foreword. J Environ Sci Health Part A-Toxic/Hazardous Subst Environ Eng 42:1191–1191

Gron C, Oomen A, Weyand E, Wittsiepe J (2007) Bioaccessibility of PAH from Danish soils. J Environ Sci Health Part A 42:1233–1239

Hamel SC, Buckley B, Lioy PJ (1998) Bioaccessibility of metals in soils for different liquid to solid ratios in synthetic gastric fluid. Environ Sci Technol 32:358–362

Hamel SC, Ellickson KM, Lioy PJ (1999) The estimation of the bioaccessibility of heavy metals in soils using artificial biofluids by two novel methods: mass-balance and soil recapture. Sci Total Environ 244:273–283

Hansen JB, Oomen AG, Edelgaard I, Gron C (2007) Oral bioaccessibility and leaching: tests for soil risk assessment. Eng Life Sci 7:170–176

Health Canada (2004) Part III: guidance on peer review of human health risk assessment, September 2004

Jimoh M, Frenzel W, Muller V (2005) Microanalytical flow-through method for assessment of the bioavailability of toxic metals in environmental samples. Anal Bioanal Chem 381:438–444

Jondreville C, Revy PS (2003) An update on use of organic minerals in swine nutrition. In: Proceedings of conference de Nutrition de l'Est du Canada, Association de Nutrition Animale du Canada, 8–9 May 2003, Quе́bec City, Quе́bec, Canada, pp 143–159

Juhasz AL, Smith E, Weber J, Naidu R, Rees M, Rofe A, Kuchel T, Sansom L (2008) Effect of soil ageing on in vivo arsenic bioavailability in two dissimilar soils. Chemosphere 71: 2180–2186

Juhasz AL, Smith E, Weber J, Rees M, Rofe A, Kuchel T, Sansom L, Naidu R (2007) in vitro assessment of arsenic bioaccessibility in contaminated (anthropogenic and geogenic) soils. Chemosphere 69:69–78

Juhasz AL, Smith E, Weber J, Rees M, Rofe A, Kuchel T, Sansom L, Naidu R (2007) Comparison of in vivo and in vitro methodologies for the assessment of arsenic bioavailability in contaminated soils. Chemosphere 69(6):961–966

Juhasz AL, Smith E, Weber J, Naidu R, Rees M, Rofe A, Kuchel T, Sansom L (2008) Effect of soil ageing on in vivo arsenic bioavailability in two dissimilar soils. Chemosphere 71:2180–2186

Kelleher AM (1999) Assessment of lead and arsenic bioavailability in surface soils in the Cardiff area. Cardiff University

Kelley ME, Brauning SE, Schoof R, Ruby MV (2002) Assessing oral bioavailability of metals in soil. Batelle, Columbus Richland

Kim DW, Kim KY, Choi BS, Youn P, Ryu DY, Klaasen CD, Park JD (2007) Regulation of metal transporters by dietary iron, and the relationship between body iron levels and cadmium uptake. Arch Toxicol 81:327–334

Koch I, Duso A, Haug C, Miskelly C, Sommerville M, Smith P, Reimer KJ (2005) Distinguishing between naturally and anthropogenically elevated arsenic at an abandoned arctic military site. Environ Foren 6:335–344

Laird, BD, Van de Wiele TR, Corriveau MC, Jamieson HE, Parsons MB, Verstraete W, Siciliano SD (2007) Gastrointestinal microbes increase arsenic bioaccessibility of ingested mine tailings using the simulator of the human intestinal microbial ecosystem. Environ Sci Technol 41: 5542–5547

Lee S (2006) Geochemistry and partitioning of trace metals in paddy soils affected by metal mine tailings in Korea. Geoderma 135:26–37

Lin Z, Puls RW (2000) Adsorption, desorption and oxidation of arsenic affected by clay minerals and aging process. Environ Geol 39:753–759

Lin ZX, Harsbo K, Ahlgren M, Qvarfort U (1998) The source and fate of Pb in contaminated soils at the urban area of Falun in central Sweden. Sci Total Environ 209:47–58

Lindsay WL (1991) Iron-oxide solubilization by organic-matter and its effect on iron availability. Plant Soil 130:27–34

Littell RC, Henry PR, Lewis AJ, Ammerman CB (1997) Estimation of relative bioavailability of nutrients using SAS procedures. J Animal Sci 75:2672–2683

Liu RQ, Zhao DY (2007a) In situ immobilization of Cu(II) in soils using a new class of iron phosphate nanoparticles. Chemosphere 68:1867–1876

Liu RQ, Zhao DY (2007b) Reducing leachability and bioaccessibilty of lead in soils using a new class of stabilized iron phosphate nanoparticles. Water Res 41:2491–2502

Ljung K, Oomen A, Duits M, Selinus O, Berglund M (2007) Bioaccessibility of metals in urban playground soils. J Environ Sci Health Part A 42:1241–1250

Lombi E, Wenzel WW, Sletten RS (1999) Arsenic adsorption by soils and iron-oxide-coated sand: kinetics and reversibility. J Plant Nutr Soil Sci-Zeitschrift Fur Pflanzenernahrung Und Bodenkunde 162:451–456

Luo L, Zhang SZ, Shan XQ, Zhu YG (2006) Effects of oxalate and humic acid on arsenate sorption by and desorption from a Chinese red soil. Water Air Soil Pollut 176:269–283

Maddaloni M, Lolacono N, Manton W, Blum C, Drexler J, Graziano J (1998) Bioavailability of soilborne lead in adults, by stable isotope dilution. Environ Health Perspect 106:1589–1594

Madrid F, Biasioli M, Ajmone-Marsan F (2008) Availability and bioaccessibility of metals in fine particles of some urban soils. Archiv Environ Contamin Toxicol 55:21–32

Manceau A, Lanson B, Schlegel ML, Harge JC, Musso M, Eybert-Berard L, Hazemann JL, Chateigner D, Lamble GM (2000) Quantitative Zn speciation in smelter-contaminated soils by EXAFS spectroscopy. Am J Sci 300:289–343

Manning BA, Goldberg S (1996) Modeling arsenate competitive adsorption on kaolinite, montmorillonite and illite. Clays Clay Min 44:609–623

Marschner B, Welge P, Hack A, Wittsiepe J, Wilhelm M (2006) Comparison of soil Pb in vitro bioaccessibility and in vivo bioavailability with Pb pools from a sequential soil extraction. Environ Sci Technol 40:2812–2818

Matera V, Le Hecho I, Laboudigue A, Thomas P, Tellier S, Astruc M (2003) A methodological approach for the identification of arsenic bearing phases in polluted soils. Environ Pollut 126:51–64

Miller ER, Ullrey DE (1987) The pig as a model for human nutrition. Annu Rev Nutr 7: 361–382

Morrison AL, Gulson BL (2007) Preliminary findings of chemistry and bioaccessibility in base metal smelter slags. Sci Total Environ 382:30–42

Mosely RA, Barnett MO, Stewart MA, Mehlhorn TL, Jardine PM, Ginder-Vogel M, Fenclorf S (2008) Decreasing Lead Bioaccessibility in Industrial and Firing Range Soils with Phosphate-Based Amendments. J Environ Qual 37:2116–2124

Moughan PJ, Birtles MJ, Cranwell PD, Smith WC, Pedraza M (1992) The piglet as a model animal for studying aspects of digestion and absorption in milk-fed human infants. World Rev Nutr Dietetics 67:40–113

Moughan PJ, Cranwell PD, Darragy AJ, Rowan AM (1994) The domestic pig as a model animal for studying digestion in humans. In: Souffrant WB, Hagemeister H (eds) Proceedings of the vth international symposium on digestive physiology in pigs, vols 1, 2, pp 389–396. EAAP European Association for Animal Production Publication

Nathanail CP (2009) Professional practice note: reviewing human health risk assessment reports invoking contaminant oral bioavailability measurements or estimates. Chartered Institute for Environment and Health. http://www.cieh.org/library/Policy/Environmental_protection/Contaminated_land/Standing_Conference_on_Contaminated_Land/CIEH_PPN_Bioavailability_Final_June09.pdf. Accessed 17 August 2010

Nathanail CP, Smith R (2007) Incorporating bioaccessibility in detailed quantitative human health risk assessments. J Environ Sci Health, Part A 42:1193–1202

Nathanail CP, McCaffrey C, Haynes D (2005b) Assessing exposure to pedogenic arsenic contamination at a dwelling in Northamptonshire, UK: a case study. Soil Use Manage 21:508–517

Nathanail CPP, McCaffrey C, Earl N, Foster ND, Gillett AG, Ogden R (2005a) A deterministic method for deriving site-specific human health assessment criteria for contaminants in soil. Hum Ecol Risk Assess 11:389–410

Nathanail CP, Mccaffrey C, Ogden, R, Foster N, Haynes D (2004) Uptake of arsenic by vegetables for human consumption: a study of wellingborough allotment plots. Land Contamin Reclamation 12

Nico PS, Ruby MV, Lowney YW, Holm SE (2006) Chemical speciation and bioaccessibility of arsenic and chromium in chromated copper arsenate-treated wood and soils. Environ Sci Technol 40:402–408

ODPM (2004) PPS23 Annex 2: development on land affected by contamination. PPS23. ODPM, London

Oomen AG (2000) Determination of oral bioavailability of soil-borne contaminants. University of Utrecht. PhD Thesis

Oomen AG, Hack A, Minekus M, Zeijdner E, Cornelis C, Schoeters G, Verstraete W, Van de Wiele T, Wragg J, Rompelberg CJM, Sips A, Van Wijnen JH (2002) Comparison of five in vitro digestion models to study the bioaccessibility of soil contaminants. Environ Sci Technol 36:3326–3334

Oomen AG, Rompelberg CJM, Bruil MA, Dobbe CJG, Pereboom D, Sips A (2003) Development of an in vitro digestion model for estimating the bioaccessibility of soil contaminants. Archiv Environ Contamin Toxicol 44:281–287

Palumbo-Roe B, Cave MR, Klinck BA, Wragg J, Taylor H, O'Donnell K, Shaw RA (2005) Bioaccessibility of arsenic in soils developed over Jurassic ironstones in eastern England. Environ Geochem Health 27:121–130

Palumbo-Roe B, Klinck B (2007) Bioaccessibility of arsenic in mine waste-contaminated soils: a case study from an abandoned arsenic mine in SW England (UK). J Environ Sci Health Part A 42:1251–1261

Park JD, Cherrington J, Klaasen CD (2002) Intestinal absorption of cadmium is associated with divalent metal transporter 1 in rats. Toxicol Sci 68:288–294

Paustenbach DJ (2000) The practice of exposure assessment: a state-of-the-art review. Reprinted from Principles and methods of toxicology, 4th edn, 2001. J Toxicol Environ Health-Part B-Crit Rev 3:179–291

Peak D, Saha UK, Huang PM (2006) Selenite adsorption mechanisms on pure and coated montmorillonite: an EXAFS and XANES spectroscopic study. Soil Sci Soc Am J 70: 192–203

Poggio L, Vrscaj B, Schulin R, Hepperle E, Marsan FA (2009) Metals pollution and human bioaccessibility of topsoils in Grugliasco (Italy). Environ Pollut 157:680–689

Pouschat P, Zagury GJ (2006) in vitro gastrointestinal bioavailability of arsenic in soils collected near CCA-treated utility poles. Environ Sci Technol 40:4317–4323

Rasmussen PE, Beauchemin S, Nugent M, Dugandzic R, Lanouette M, Chenier M (2008) Influence of matrix composition on the bioaccessibility of copper, zinc and nickel in urban residential dust and soil. Human Ecol Risk Assess 14:351–371

Reeder RJ, Schoonen MAA, Lanzirotti A (2006) Metal speciation and its role in bioaccessibility and bioavailability. Med Mineraol Geochem. Rev Mineral Geochem 64:59–113 (Chantilly: Mineralogical Soc Am) ISBN 1529-6466

Ren HM, Wang JD, Zhang XL (2006) Assessment of soil lead exposure in children in Shenyang China. Environ Pollut 144:327–335

Roberts SM, John W, Munson JW, Lowney YW, Ruby MV (2007a) Relative oral bioavailability of arsenic from contaminated soils measured in the cynomolgus monkey. Toxicol Sci 95: 281–288

Roberts SM, Munson JW, Lowney YW, Ruby MV (2007b) Relative oral bioavailability of arsenic from contaminated soils measured in the cynomolgus monkey. Toxicol Sci 95:281–288

Rodriguez RR, Basta NT, Casteel S, Pace L (1999) An in vitro gastrointestinal method to estimate bioavailable arsenic in contaminated soils and solid media. Environ Sci Technol 33:642–649

Rodriguez RR, Basta NT, Casteel SW, Armstrong FP, Ward DC (2003) Chemical extraction methods to assess bioavailable arsenic in soil and solid media. J Environ Qual 32(3): 876–884

Romero FM, Villalobos M, Aguirre R, Gutierrez ME (2008) Solid-phase control on the lead bioaccessibility in smelter-impacted soils. Archiv Environ Contamin Toxicol 55:566–575

Rowan A, Moughan M, Wilson PJ, Maher MN, Tasman K, Jones C (1994) Comparison of the ileal and feacal digestibility of dietary amino acids in adult humans and evaluation of the pig as a model animal for digestion studies in man. Br J Nutr 71:29–42

Rose AW, Hawkes HE, Webb JS (1990) Geochemical methods in mineral exploration. Geochemistry in mineral exploration. Academic, London, pp 535–548

Ruby MV (2004) Bioavailability of soil-borne chemicals: abiotic assessment tools. Human Ecol Risk Assess 10:647–656

Ruby MV, Davis A, Link TE, Schoof R, Chaney RL, Freeman GB, Bergstrom P (1993) Development of an in vitro screening test to evaluate the in vivo bioaccessibility of ingested mine waste lead. Environ Sci Technol 27:2870–2877

Ruby MV, Davis A, Schoof R, Eberle S, Sellstone CM (1996) Estimation of lead and arsenic bioavailability using a physiologically based extraction test. Environ Sci Technol 30: 422–430

Ruby MV, Fehling KA, Paustenbach DJ, Landenberger BD, Holsapple MP (2002) Oral bioaccessibility of dioxins/furans at low concentrations (50–350 ppt toxicity equivalent) in soil. Environ Sci Technol 35:4905–4911

Ruby MV, Schoof R, Brattin W, Goldade M, Post G, Harnois M, Mosby DE, Casteel SW, Berti W, Carpenter M, Edwards D, Cragin D, Chappell W (1999) Advances in evaluating the oral bioavailability of inorganics in soil for use in human health risk assessment. Environ Sci Technol 33:3697–3705

Sarkar D, Makris KC, Parra-Noonan MT, Datta R (2007b) Effect of soil properties on arsenic fractionation and bioaccessibility in cattle and sheep dipping vat sites. Environ Int 33:164–169

Sarkar D, Quazi S, Makris KC, Datta R, Khairom A (2007a) Arsenic bioaccessibility in a soil amended with drinking water treatment residuals in the presence of phosphorous fertiliser. Archiv Environ Contamin Toxicol 53:329–336

Schaider LA, Senn DB, Brabander DJ, McCarthy KD, Shine JP (2007) Characterization of zinc, lead, and cadmium in mine waste: implications for transport, exposure, and bioavailability. Environ Sci Technol 41:4164–4171

Schroder JL, Basta NT, Casteel SW, Evans TJ, Payton ME, Si J (2004a) Validation of the in vitro gastrointestinal (IVG) method to estimate relative bioavailable lead in contaminated soils. J Environ Qual 33:513–521

Schroder JL, Basta NT, Si JT, Casteel SW, Evans T, Payton M (2003) in vitro gastrointestinal method to estimate relative bioavailable cadmium in contaminated soil. Environ Sci Technol 37:1365–1370

Schroder JL, Basta NT, Si JT, Casteel SW, Evans T, Payton M, Si P (2004b) Validation of the in vitro gastointestinal (IVG) method to estimate relative bioavailable lead in contaminated soils. J Environ Qual 33(2):513–521

Schwab AP, Lewis K, Banks MK (2006) Lead stabilisation by phosphate ammendments in soil impacted by paint residue. J Environ Sci Health Part A 41:359–368

Selinus O. (2005) Essentials of medical geology impacts of the natural environment on public health. Elsevier Academic, London

Sharp AP (2003) Ctr1 and its role in body copper homeostasis. Int J Biochem Cell Biol 35:288–291

Siebielec G, Stuczynski T, Korzeniowska-Puculek R (2006) Metal bioavailability in long-term contaminated Tarnowskie Gory soils. Polish J Environ Studies 15:121–129

Smith E, Naidu R, Alston AM (1998) Arsenic in the soil environment: a review. Adv Agron 64:149–195

Smith E, Naidu R, Weber J, Juhasz AL (2008) The impact of sequestration on the bioaccessibilty of arsenic in long-term contaminated soil. Chemosphere 71:773–780

Stewart MA, Jardine PM, Barnett MO, Mehlhorn TL, Hyder LK, McKay LD (2003a) Influence of soil geochemical and physical properties on the sorption and bioaccessibility of chromium(III). J Environ Qual 32:129–137

Stewart MA, Jardine PM, Brandt CC, Barnett MO, Fendorf SE, McKay LD, Mehlhorn TL, Paul K (2003b) Effects of contaminant concentration, aging, and soil properties on the bioaccessibility of Cr(III) and Cr(VI) in soil. Soil Sediment Contamin 12:1–21

Somez O, Pierzynski GM (2005) Phosphorous and manganese oxides effects on soil lead bioaccessibility: PBET and TCLP. Water Air Soil Pollut 166:3–16

Subacz JL, Barnett MO, Jardine PM, Steward MA (2007) Decreasing arsenic bioaccessibility/bioavailability in soils with iron ammendments. J Environ Sci Health Part A 42:1317–1329

Sultan K (2007) Distribution of metals and arsenic in soils of Central Victoria (Creswick-Ballarat), Australia. Archiv Environ Contamin Toxicol 52:339–346

Tallkvist J, Bowlus C, Lonnerdal B (2001) DMT1 gene expression and cadmium absorption in human absportive enterocytes. Toxicol Lett 122:171–177

Tang XY, Zhu YG, Chen SB, Tang LL, Chen XP (2004) Assessment of the effectiveness of different phosphorus fertilizers to remediate Pb-contaminated soil using in vitro test. Environ Int 30: 531–537

Tang XY, Zhu YG, Cui YS, Duan J, Tang L (2006) The effect of ageing on the bioaccessibility and fractionation of cadmium in some typical soils of China. Environ Int 32:682–689

Tang X-Y, Zhu Y-G, Shan X-Q, McLaren R, Duan J (2007) The ageing effect on the bioaccessibility and fractionation of arsenic in soils from China. Chemosphere 66:1183–1190

Tilston E, Collins C, Gibson G, Cave M, Wragg J (2008) Bioaccessibility of organic pollutants in the human digestive system. Consoil 2008, Milan, Italy

Tseng C-H (2009) A review of environmental factors regulating arsenic methylation in humans. Contemporary Issues Toxicol 235:338–350

United States Environmental Protection Agency (1989) Risk assessment guidance for superfund, volume I: human health evaluation manual (Part A)

United States Environmental Protection Agency (2007a) Guidance for Evaluating the oral bioavailability of metals in soils for use in human health risk assessment, OSWER 9285.7-80

United States Environmental Protection Agency (2007b) Estimation of relative bioavailability of lead in soil and soil-like material using in vivo and in vitro methods, OSWER 9285.7-77

Van de Wiele T, Vanhaecke L, Boeckaert C, Peru K, Headley J, Verstraete W, Siciliano S (2005) Human colon microbiota transform polycyclic aromatic hydrocarbons to estrogenic metabolites. Environ Health Perspect 113:7–10

Van de Wiele TR, Oomen AG, Wragg J, Cave M, Minekus M, Hack A, Cornelis C, Rompelberg CJM, De Zwart LL, Klinck B, Van Wijnen J, Verstraete W, Sips AJAM (2007) Comparison of five in vitro digestion models to in vivo experimental results: lead bioaccessibility in the human gastrointestinal tract. J Environ Sci Health, Part A 42:1203–1211

Versantvoort CHM, Van de Kamp E, Rompelberg CJM (2004) Development and applicability of an in vitro digestion model in assessing the bioaccessibility of contaminants from food. RIVM Report 320102002, Bilthoven, Netherlands

Violante A, Pigna M (2002) Competitive sorption of arsenate and phosphate on different clay minerals and soils. Soil Sci Soc Am J 66:1788–1796

Violante A, Ricciardella M, Del Gaudio S, Pigna M (2006) Coprecipitation of arsenate with metal oxides: nature, mineralogy, and reactivity of aluminum precipitates. Environ Sci Technol 40:4961–4967

Webb M (1986) Role of metallothionein in cadmium metabolism. In: Foulkes EC (ed) Cadmium, vol 80. Springer, New York, pp 281–338

Welter E, Calmano W, Mangold S, Troger L (1999) Chemical speciation of heavy metals in soils by use of XAFS spectroscopy and electron microscopical techniques. Fresenius J Anal Chem 364:238–244

Wiettsiepe J, Schrey P, Hack A, Selenka F, Wilhelm M (2001) Comparison of different digestive tract models for estimating bioaccessibility of polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/F) from red slag ‘Kieselrot’. Int J Hygiene Environ Health 203:263–273

Wragg J (2005) A study of the relationship between arsenic bioaccessibility and its solid phase distribution in Wellingborough soils. Nottingham

Wragg J, Cave M, Nathanail P (2007) A Study of the relationship between arsenic bioaccessibility and its solid-phase distribution in soils from Wellingborough, UK. J Environ Sci Health, Part A 42:1303–1315

Wragg J, Cave M, Taylor H (2009) Inter-laboratory trial of a unified bioaccessibility procedure. British Geological Survey, OR/07/027

Yang JK, Barnett MO, Jardine PM, Basta NT, Casteel SW (2002) Adsorption, sequestration, and bioaccessibility of As(V) in soils. Environ Sci Technol 36:4562–4569

Yang JK, Barnett MO, Jardine PM, Brooks SC (2003) Factors controlling the bioaccessibility of arsenic(V) and lead(II) in soil. Soil Sediment Contamin 12:165–179

Yang JK, Barnett MO, Zhuang JL, Fendorf SE, Jardine PM (2005) Adsorption, oxidation, and bioaccessibility of As(III) in soils. Environ Sci Technol 39:7102–7110

Zagury GJ (2007) Comments on “effect of soil properties on arsenic fractionation and bioaccessibility in cattle and sheep dipping vat sites” by Sarkar D et al (2007) Environ Int 33:164–169. Environ Int 33:712–713

Chapter 8
Uptake of Metals from Soil into Vegetables

Mike J. McLaughlin, Erik Smolders, Fien Degryse, and Rene Rietra

Abstract The consumption of locally-produced vegetables by humans may be an important exposure pathway for soil contaminants in many urban settings and for agricultural land use. Hence, prediction of metal and metalloid uptake by vegetables from contaminated soils is an important part of the Human Health Risk Assessment procedure. The behaviour of metals (cadmium, chromium, cobalt, copper, mercury, molybdenum, nickel, lead and zinc) and metalloids (arsenic, boron and selenium) in contaminated soils depends to a large extent on the intrinsic charge, valence and speciation of the contaminant ion, and soil properties such as pH, redox status and contents of clay and/or organic matter. However, chemistry and behaviour of the contaminant in soil alone cannot predict soil-to-plant transfer. Root uptake, root selectivity, ion interactions, rhizosphere processes, leaf uptake from the atmosphere, and plant partitioning are important processes that ultimately govern the accumulation of metals and metalloids in edible vegetable tissues. Mechanistic models to accurately describe all these processes have not yet been developed, let alone validated under field conditions. Hence, to estimate risks by vegetable consumption, empirical models have been used to correlate concentrations of metals and metalloids in contaminated soils, soil physico-chemical characteristics, and concentrations of elements in vegetable tissues. These models should only be used within the bounds of their calibration, and often need to be re-calibrated or validated using local soil and environmental conditions on a regional or site-specific basis.

M.J. McLaughlin (✉)
CSIRO Land and Water/University of Adelaide, Glen Osmond, SA, Australia
e-mail: Mike.McLaughlin@csiro.au

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_8,

Contents

8.1 Introduction

The transfer of metal (and metalloid) contaminants from soil, through plants, to humans, may be an important exposure pathway in some urban and residential environments. This exposure pathway may also be important for agricultural land-use scenarios to manage exposure of the general population to metals and metalloids. This chapter summarises our current understanding of metal/metalloid behaviour in soils, uptake and transport of these elements by plants, and the current models and concepts used to predict exposure of humans to metals and metalloids through a food-chain pathway.

8.2 Metal and Metalloid Chemistry in Soil

In describing the chemistry of metals and metalloids in soil, distinctions need to be drawn between the behaviour of cationic and anionic elements given the importance of surface charge to the fate and behaviour of elements in soils (Sposito 1981, 1989). Most topsoils have a net negative surface charge, with exceptions being soils rich in iron, aluminium or manganese oxides and depleted in organic matter and phosphorus (P) or sulphur (S). These latter soils, which may have a significant net positive charge, are found in tropical regions, particularly in subsoils (Wong and Wittwer 2009). Soils with net negative charge retain cationic metals more strongly, soils with net positive charge will retain anionic metals more strongly. Hence we will discuss the behaviour of cationic and anionic metals separately. A further key property of metals which controls environmental fate is oxidation state, as this may affect charge e.g. chromium (Cr) – Cr^{3+} present as this trivalent cation in most soils, and Cr^{6+}, present as an oxyanions as $CrO_4{}^{2-}$ or $HCrO_4{}^{-}$.

8.2.1 Cationic Metals

Elements in this group include cadmium (Cd), cobalt (Co), chromium (Cr), copper (Cu), mercury (Hg), nickel (Ni), lead (Pb) and zinc (Zn). These elements may have different valence states, but are generally present in cationic form in soil, either as free metals in solution (M^{n+}), metals complexed to inorganic or organic ligands (ML) or as solid phases (e.g. precipitated as minerals, or sorbed to negatively charged surfaces) (McLaughlin 2002 – Table 8.1). The distribution coefficient (K_d) is a measure of the relative distribution of elements between the solid phase and pore water (water held in pore space between soil particles), and is an important property to consider in assessing the potential availability of metals to plants:

$$K_d = M_{\text{solid}} / M_{\text{pore water}} \tag{8.1}$$

where M_{solid} = concentration in the solid phase ($\text{mg}\cdot\text{kg}_{\text{dw}}{}^{-1}$) and $M_{\text{pore water}}$ is the concentration in the pore water ($\text{mg}\cdot\text{L}^{-1}$).

K_d values for cationic metals vary widely across soils, and are also dependent on metal properties such as valence state, e.g. Co^{3+} has much higher K_d values in soils than Co^{2+}. Metal partitioning in soils can be due to both sorption/desorption

Table 8.1 Physical and chemical properties of the cationic metals

Element	Symbol	Valence	Dominant species in soil	Dominant species in soil solution	
				pH 3.5–6.0	pH 6.0–8.5
Cadmium	Cd	2	Cd^{2+}	Cd^{2+}, $CdCl^{+}$, $CdSO_4{}^{0}$	Cd^{2+}, $CdCl^{+}$, $CdSO_4{}^{0}$
Chromium	Cr	2,3,6	Cr^{3+}, $CrO_4{}^{2-}$	Cr^{3+}, $CrOH^{2+}$	$Cr(OH)_4{}^{-}$
Cobalt	Co	2,3	Co^{2+}	–	–
Copper	Cu	1,2	Cu^{2+}	Cu^{2+}, Cu-org.	Cu-hydroxy species, $CuCO_3{}^{0}$, Cu-org.
Lead	Pb	2,4	Pb^{2+}	Pb^{2+}, $PbSO_4{}^{0}$, Pb-org.	Pb-hydroxy and carbonate species, Pb-org.
Mercury	Hg	1,2	Hg^{2+}, $(CH_3)_2Hg$	–	–
Nickel	Ni	2,3	Ni^{2+}	Ni^{2+}, $NiSO_4{}^{0}$, Ni-org.	Ni^{2+}, $NiHCO_3{}^{+}$, $NiCO_3$
Zinc	Zn	2	Zn^{2+}	Zn^{2+}, $ZnSO_4{}^{0}$, Zn-org.	Zn^{2+}, Zn-hydroxy and carbonate species, Zn-org.

Adapted from McLaughlin (2002)

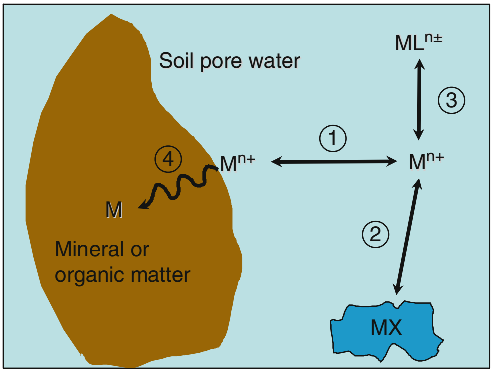


Fig. 8.1 Relationship between cationic metals in discrete mineral phases (MX), free cationic metal ions in pore water (M^{n+}), complexed metals in pore water (ML), metals sorbed to soil charged surfaces, and metals occluded within minerals or organic matter in soil. Reaction ① represents sorption/desorption reactions and is described by a sorption coefficient, reaction ② represents precipitation/dissolution reactions and is described by a solubility product, reaction ③ represents a solution complexation reaction and is described by an association constant, and reaction ④ is generally termed an 'ageing' reaction whereby sorbed metals become less available with time (McLaughlin 2001b)

reactions, as well as precipitation/dissolution reactions (Fig. 8.1). This obviously depends on the degree of and type of soil contamination. Soils highly contaminated by soluble metal sources are more likely to have distinguishable metal precipitates, as are those soils contaminated by sparingly soluble materials. Sorption/desorption processes are more likely to control metal partitioning in soils contaminated by lower concentrations of soluble contaminants, or in soils contaminated by highly soluble solid materials. The impact of both reactions combined is represented by the K_d.

There have been numerous studies of metal partitioning in soils and how K_d values are affected by soil mineralogy, particle size, pH, salinity, redox status, metal loading, etc. and the reader is referred to several review articles that summarise these findings (Buchter et al. 1989; Degryse et al. 2009; McBride 1989; Sauvé et al. 2000). The ranges of K_d values found for various metals are shown in Table 8.2.

The most important factor found to influence the dissolution of cationic metal precipitates, and the release of cationic metals by negatively charged soil surfaces, is soil pH. Hence higher concentrations of cationic metals are found in the pore waters of acidic soils. Moreover, raising soil pH by liming will significantly reduce cationic metal concentrations in soil pore waters. Soil pH is therefore a master variable in controlling the distribution of cationic metals between the solid phase and the pore water, and hence their availability to plants. Soil particle size and mineralogy also influence K_d values, with sandy soils or soils having low concentrations of high surface area minerals having low K_d values. Sandy acidic soils are therefore most at risk of having high concentrations of cationic metals in soil pore water, and hence a risk of plant uptake of these contaminants. Due to the inverse relationship

Table 8.2 Range of K_d values for cationic metals and anionic metals/metalloids in soils

Element	K_d (mean)	SD	CV	Median	Min	Max	log 10 K_d	N
Cationic metals								
Cd	2,869	12,246	4.27	390	0.44	1,92,000	3.46	830
Cu	4,799	9,875	2.06	2,120	6.8	82,850	3.68	452
Hg	8,946	5,641	0.63	7,500	4286	16,500	3.95	4
Ni	16,761	45,350	2.71	2,333	8.9	2,56,842	4.22	139
Pb	1,71,214	3,04,089	1.78	1,02,410	60.56	23,04,762	5.23	204
Zn	11,615	30,693	2.64	1,731	1.4	3,20,000	4.07	302
Anionic metals/metalloids								
As	13,119	65,086	4.96	1,825	1.6	5,30,000	4.12	66
B	160	96	0.60	136	61	389	2.20	12
Mo	36	19	0.52	38	14	52	1.55	4
Se	43,937	1,19,534	2.72	15	1.6	6,00,000	4.64	63

Sauvé et al. (2000)

between metal loading to soil and metal K_d (Hendrickson and Corey 1981), highly contaminated soils are likely to have lower metal retention than soils contaminated by lower metal loads.

While K_d values are normally determined experimentally in the laboratory, for many metals and metalloids approximate K_d values can be estimated from empirical relationships derived using total metal concentrations in soil, and the main factors affecting partitioning, usually pH, and soil clay/oxide or organic matter content (Anderson and Christensen 1988; Buchter et al. 1989; Sauvé et al. 2000). These models do not require detailed input data (as do more mechanistic models of metal partitioning). Therefore, they can be a useful first screening tool in first tier Risk Assessments to assess *potential* metal availability at a contaminated site, using very simple soil analytical data – total metal concentrations and pH and/or clay/organic matter content.

Cationic metals may also react with soluble natural or anthropogenic ligands (e.g. chloride, dissolved organic matter or synthetic chelates such as ethylenediaminetetraacetate – EDTA) to form solution complexes in soil pore waters (Fig. 8.1). This may reduce or even reverse the charge on the cationic metal and hence markedly reduce the K_d value, increasing metal availability, but not necessarily increasing metal uptake as this is governed by the free metal ion concentration in the pore water and other factors (see Section 8.4.1).

8.2.2 Anionic Metals/Metalloids

Several of the metals and metalloids may be found in anionic form in soils, due to their ability to combine with oxygen to form oxyanions (Table 8.3). These ions may not be held as strongly by soil due to the net negative charge in most soils,

Table 8.3 Physical and chemical properties of the anionic metals/metalloids

Element	Symbol	Valence	Dominant species in soil	Dominant species in soil solution	
				pH 3.5–6.0	pH 6.0–8.5
Arsenic	As	−3,0,3,5	As^{3+}, As^{5+}	$H_2AsO_4^-$, H_3AsO_3	$HAsO_4^{2-}$
Boron	B	3	B^{3+}	$B(OH)_3$	$B(OH)_3$, $B(OH)_4^-$
Chromium	Cr	2,3,6	Cr^{3+},CrO_4^{2-}	Cr^{3+}, $CrOH^{2+}$	$Cr(OH)_4^-$
Molybdenum	Mo	2,3,4,5,6	Mo^{6+}	$HMoO_4^-$, MoO_4^{2-}, H_2MoO_4	MoO_4^{2-}
Selenium	Se	0,2,3,4,6	Se^0, Se^{4+}, Se^{6+}	SeO_4^{2-}, $HSeO_3^-$	SeO_4^{2-}, SeO_3^{2-}

and hence they may have relatively low K_d values compared to the cationic metals. Hence these elements are much more readily available for plant uptake from soil compared to cationic metals with high K_d values (e.g. lead). Exceptions to this rule are those anions that can form strong bonds through ligand-exchange with soil minerals, e.g. phosphate, fluoride, but there are few metals or metalloids with this property. Arsenate is one of the few metalloids that may associate with soil minerals through ligand exchange reactions, and hence have stronger binding than would be suggested by consideration of charge alone.

In contrast to cationic metals, the oxyanions have higher K_d values in acidic soils compared to alkaline soils, and different redox species of the same metal/metalloid may have widely differing K_d values (Bartlett and James 1988; Bowell 1994; Goldberg 1997; Goldberg and Forster 1998; Nakamaru et al. 2005). Arsenic is a good example, where As^{5+}, usually present in soils as the arsenate ion $H_xAsO_4^{-(3-x)}$ generally has higher K_d values in soil than As^{3+}, present as H_3AsO_3 (Bowell 1994).

Of the metals normally found at contaminated sites, chromium perhaps is unique in that a valency change causes charge reversal, and this markedly affects not only chromium partitioning (Bartlett and James 1988), but also toxicity (McGrath 1982), with the Cr^{6+} species being more available and toxic compared to Cr^{3+}.

8.2.3 Effects of Soil Redox

As outlined above, changes in soil oxidation/reduction potential have the capability to alter the valence state of metal/metalloids present in soil. However, soil redox has a more important effect on metal chemistry in that the major components of soil which are active in metal/metalloid retention, Fe and Mn oxyhydroxides, are also redox sensitive principally through the Fe^{2+}/Fe^{3+} and Mn^{2+}/Mn^{4+} couples. Both Fe^{2+} and Mn^{2+} species are much more soluble than their oxidised species, so that reduction of soil, e.g. due to waterlogging, will cause reductive dissolution of Fe and Mn oxides, and often destruction of the surfaces active in the retention of many metals and metalloid ions (Plekhanova 2007). Thus, water logging of soil will lead to the release of elements strongly retained by these surfaces, e.g. arsenic concentrations are much higher in pore waters of waterlogged soils than aerobic soils (Bowell 1994; Marin et al. 1992).

8.3 Plant Acquisition of Metals and Metalloids from Soil

8.3.1 Root Uptake Pathway

In this section, we will elaborate on the processes by which metals and metalloids are taken up from soil by plant roots and how these elements are subsequently distributed with the plant tissues.

8.3.1.1 Speciation and Ion Uptake Rate

Plant roots absorb metals/metalloids via the root bathing pore water. The metals/metalloids commonly enter the plant as ions (e.g. Cd^{2+}, $H_2AsO_4^-$) *via* ion channels or carriers that have the capacity to concentrate the elements from solution. *Non-essential elements* enter the plants using the uptake systems of nutrients that resemble the contaminants in terms of charge and ionic radius. Passive uptake of the elements through water uptake rarely explains the observed uptake of several nutrients (Barber 1995) and the same is true for contaminants such as cadmium and lead. The uptake rate generally increases with increasing concentration in pore water. Short-term ion uptake studies with roots demonstrate that uptake of contaminants follows a concentration-dependent pattern that is similar to enzyme kinetics. In a similar, but not identical, fashion it is observed that *plant tissue concentrations rise as soil concentrations rise*. Such patterns are important in Risk Assessment where the concept is to identify tolerable soil concentrations at which plant concentrations are below target values. Figure 8.2a summarises the general patterns for non-essential and essential contaminants. The tissue concentrations of non-essential elements such as arsenic, cadmium, mercury, and lead rise almost proportionally to their concentrations in soil at low concentrations. This pattern is the basis of using the *BioConcentration Factor (BCF) concept* in Risk Assessment, i.e. a constant plant tissue:soil concentration relationship. As soil concentrations rise, however, there are saturating processes and tissue concentration levels off, resulting in BCF values that are lower than that at low concentrations. The saturating processes are root uptake or translocation processes or potential feedback mechanisms under toxic conditions. For essential elements, such as the micronutrients boron (B), copper (Cu), manganese (Mn), molybdenum (Mo), and zinc (Zn), the pattern is distinctly different. Tissue concentrations are maintained within narrow limits at widely different external concentrations through homeostatic control mechanisms on ion uptake and translocation. At elevated supply, however, the control mechanisms break down

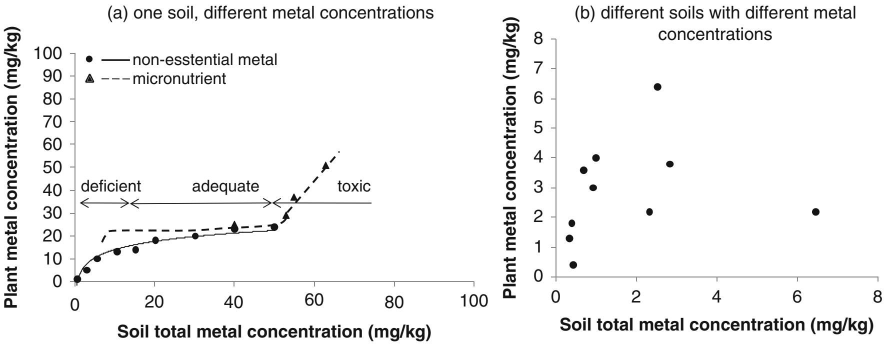


Fig. 8.2 (**a**) Generalized concentration dependent uptake for metals:homeostatic processes maintain internal concentrations of essential metals at adequate supply, whereas these processes break down at phytotoxic concentrations. Uptake of non-essential metals increases proportionally to soil concentrations up to a level where saturation occurs. (**b**) the concentration dependent uptake of non-essential metals is masked by differences in soil metal bioavailability in soils with different soil properties. Data in (**a**): zinc (micronutrient) uptake by corn shoot from Zn^{2+} salt spiked soil (data of Maclean 1974) and cadmium (non-essential) uptake in soybean shoot from Cd^{2+} spiked soil (Haghiri 1973); in (**b**): cadmium uptake in wheat seedlings (Smolders et al. 1999)

and tissue concentrations readily increase with soil concentrations and this increase is generally associated with the onset of *phytotoxic conditions*. The homeostatic mechanisms generally act stronger on shoot than on root concentrations.

The soil-plant concentration relationships depicted in Fig. 8.2 are only observed in experiments where soils are enriched with soluble forms of the metal or metalloid. These patterns are termed the ‘metal or metalloid salt linear response relationships’ (Brown et al. 1998). In the environment, no such clear pattern is found, because the different bioavailabilities of the metal or metalloid in soil obscure the underlying patterns. Figure 8.2b illustrates that the concentration dependency can even be completely invisible when merging data from various soils. Variable bioavailability is related to differences in metal or metalloid speciation, interionic effects on ion uptake from pore water and indirect effects of soil properties on translocation within the plant.

The general paradigm in metal and metalloid uptake in soil is that roots absorb elements through pore water and the concentration of the dissolved elements affects the uptake rate. Dissolved metals and metalloids can be present in different forms (‘species’) and it is known that root uptake rate changes with the type of species present. Numerous experiments have shown that uptake of cationic metals from pore water generally decreases as the metal is complexed by chelating agents. Figure 8.3a illustrates this for the uptake of cadmium in a plant grown in solution culture (hydroponics). The addition of the cadmium chelator reduces the uptake of cadmium at constant cadmium concentrations, illustrating that free ion is the preferred species. This observation can be explained by the fact that ions are absorbed through their specific uptake systems while complexes are too large to pass the root cell membranes. This model is the basis of the *free ion activity model* (FIAM) that relates

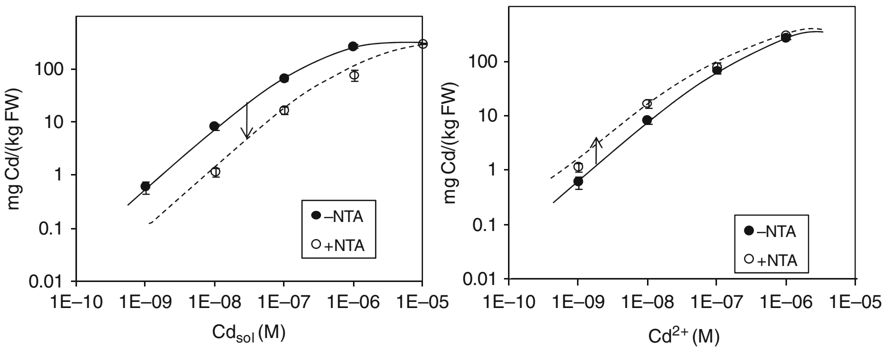


Fig. 8.3 Speciation affects shoot cadmium concentration in *Thlaspi caerulescens* grown in nutrient solution. Left: the addition of the Cd^{2+} chelator nitrilotriacetic acid (NTA) decreases cadmium uptake when compared at equal soluble cadmium concentration. These data illustrate that the free ion is preferred compared to complexed cadmium. Right: same uptake data as left but plotted to the predicted free ion activity in solution, illustrating that the chelator addition increases the availability of the free metal ion, i.e. the Free Ion Activity uptake Model (FIAM) is not fully valid. Note that the curves merge at large free ion activity suggesting that the deviation of FIAM at lower activities are not due to partial uptake of Cd-NTA but due to the mechanism of buffering Cd^{2+} in the unstirred layer adjacent to roots (see text for argumentation; unpublished data from F. Degryse)

root uptake rate to the free metal ion activity in solution (Parker et al. 1995). For that reason, experimental methods have been developed to measure the free metal ion activity in soil. The FIAM has been contested, however, and it is now clear that the free metal ion is generally the preferred species, but that the complexed (and adsorbed) species contribute to availability as well, depending on timescales considered. For example, it is observed that metal uptake increases in soil when the concentration of a complexed metal increases at constant free metal activity (e.g. Smolders and McLaughlin 1996a, b). Such is also illustrated in Fig. 8.3b that presents the same data as in Fig. 8.3a but now plotted in a free metal ion activity basis. Here, it is shown that the complexed metals have an ‘apparent availability’ in nutrient solution. Experimental evidence has now clarified that this apparent availability is related to the ability of a metal complex to buffer the free metal ion uptake at the root surface (Degryse et al. 2006a). In soils and even in stirred nutrient solution, there is a zone adjacent to roots where the rapid intake of the metal or metalloid ion is not readily replenished by the flow of water to the roots. As a result, the free ion activity decreases near the root surface (Fig. 8.4). With increasing concentrations of soluble complexes and at constant free ion activity, such depletions are less pronounced and bioavailability of the metal is enhanced. Effectively, this means that the apparent availability of the metal complex is not by direct uptake of the complex, but a consequence of the lack of a fully mixed system in the root environment. This concept now also explains why root exudates enhance both solubility as well as bioavailability of the metals copper and zinc, despite the fact that the exudates do not change the free metal ion activity in soil (Degryse et al. 2008). There is also evidence that intact metal complexes can be absorbed (Collins et al. 2001, Tandy

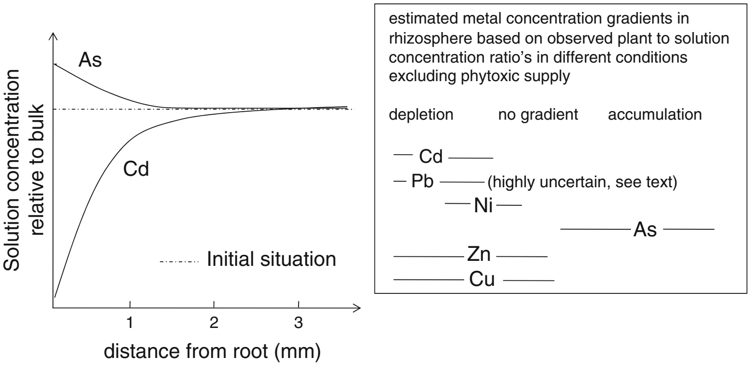


Fig. 8.4 Generalised scheme of dissolved metal or metalloid concentration gradients in the rhizosphere as a result of the ion uptake by roots, mass flow of water and diffusion to the roots. Depleted conditions occur for metals that are taken up with high absorption power such as cadmium. Rhizosphere pore water cannot be sampled equivocally, therefore it becomes more difficult to interpret metal uptake from bulk soil solution data when concentration gradients are more pronounced. The range of estimated gradients is based on an assumed averaged water use of 200 l/g dry matter and on solution culture data (e.g. Checkai et al. 1987; Degryse et al. 2006b; Weng et al. 2003) or field based plant-soil solution concentration ratios (Chen et al. 2009). The estimates for lead are most uncertain due to the lack of solution culture data at background lead exposure and since field data are confounded by atmospheric contribution of lead to aerial plant parts

et al. 2006a). For example, synthetic chelators added to soil increase shoot lead concentrations (Tandy et al. 2006a). The suspected mechanism here is *passive (non-selective) uptake* of metal-complexes through breaks in the endodermis of roots, and transfer to above ground parts in the xylem (McLaughlin et al. 1997; Tandy et al. 2006b). In that scheme, uptake increases linearly with pore water concentration and has no saturable component.

Recent large-scale surveys of plant metal concentrations and speciation of metals in the associated soils demonstrated that pore water concentrations or free ion activities of the metals did *not* explain the crop concentrations (Chen et al. 2009; Hough et al. 2005). This observation does not invalidate the general concept that pore water concentrations and free ions are the directly available forms of these elements. Two processes may be invoked here that complicate the analysis in practice, i.e. *rhizosphere processes* that alter the pore water composition compared to the solution that can be sampled, and *ion competition* effects which affect the uptake rate of the free ion.

8.3.1.2 Rhizosphere Processes

The *rhizosphere* is the soil environment influenced by the roots and the rhizosphere conditions effectively control the supply of contaminants to plant roots. Unfortunately, the rhizosphere cannot be sampled unequivocally and this is a known drawback in soil bioavailability research. The physicochemical processes that alter metal speciation in the rhizosphere are the gradients in pH, soluble organic matter, depletions or accumulation of ions and redox gradients. We refer the interested reader to reviews dedicated to this topic (McLaughlin et al. 1998) and only explore the causes of concentration gradients, since this has practical consequences as shown below. Concentration gradients in the rhizosphere are the result of the balance between the ion uptake rate and resupply by the uptake of water. A simple calculation allows estimating if the contaminant is either accumulated or depleted in the rhizosphere: a plant typically transpires about 200 L of water per kg dry weight produced (Barber 1995). The product of the transpiration and the concentration of the contaminant in the pore water is termed the *mass flow* of elements to the plant root. If mass flow matches uptake perfectly, then the ratio of the contaminant concentration in the plant (mg/kg$_{dw}$) to that in *pore water* (mg/L) should be 200. Larger ratios mean that mass flow is not sufficient to match the rate of element uptake by the root, and concentrations of that element will have been depleted in the rhizosphere. The concentration depletion is followed by a diffusion flux towards the roots and this flux can be several fold larger than mass flow. Conversely, if concentration ratios are lower than 200, then mass flow (induced by transpiration) has exceeded root uptake and, consequently, the contaminant may have accumulated around roots. Field-based data for several plants and soils show that cadmium and lead concentration ratios exceed this threshold by 1–2 orders of magnitude, while arsenic concentrations ratios are typically 1 order of magnitude below that (Chen et al. 2009). This means that cadmium and lead are, on average, depleted in the rhizosphere, while arsenic generally accumulates around the roots. The differences are related to differences in the so-called root absorption power (uptake rate per unit concentration), for example a low value for arsenic uptake per unit time from the pore water due to the strong competition with phosphate ions. Similar data for other metals and metalloids allows ranking as given in Fig. 8.4. The practical consequence of these gradients is that the causal relationship between pore water concentrations and tissue concentrations, observed in stirred solution, may not be detectable in soils anymore because we fail to measure the pore water concentrations in the rhizosphere. The rhizosphere conditions can be estimated by modelling the diffusion and mass flow (Barber 1995). If the element is depleted in the rhizosphere, then it is replenished by either solid or liquid complexed forms. Practically, this means that a fraction of the complexed ions is also part of the directly available forms in soil, provided that the dissociation rate is sufficiently rapid. For that reason, bioavailability is always a complex function of both the activity in the pore water and a fraction of the labile bound forms. It is also logical that assessments of diffusive fluxes (i.e. fluxes under conditions of zero-sink) correlates well with the uptake, provided that the metal or metalloid is indeed depleted in the rhizosphere (Nolan et al. 2005).

8.3.1.3 Ion Competition Effects for Metal and Metalloid Uptake

Ion uptake is furthermore affected by *interionic effects*, i.e. the uptake rate of the ion decreases or increases as the concentration of an ion competing with the same uptake site is increased or decreased, respectively. One of the most striking examples of this is the effect of pH on uptake of cationic metals. For example, different solution culture studies have shown that metal uptake increases when H^+ activity is lowered (pH increases). For example, Weng et al. (2003) showed that nickel uptake in oats increased by about a factor of 3 per unit pH increase between pH 4–6, at constant Ni^{2+} activity. Concentrations of cadmium in soybean shoots was shown to increase by a factor of 1.9 between pH 5–7 at constant Cd^{2+} activity (Smolders and Helmke unpublished), while shoot cadmium increased markedly larger, with factors of 4–13 in unbuffered nutrient solution between pH 5–7 for ryegrass, lettuce, cockfoot and watercress (Hatch et al. 1988). Numerous other interaction effects in plants have been identified for metal and metalloid uptake. Without attempting to be complete, we cite those that are relevant for contaminated site Risk Assessment, i.e. Ca^{2+}:Cd^{2+} (Tyler and McBride 1982), Zn^{2+}:Cd^{2+} (McKenna et al. 1993), H^+:Cu^{2+} (Chen and Allen 2001), $H_2PO_4^-$:$H_2AsO_4^-$ (Khattak et al. 1991) and SO_4^{2-}:SeO_4^{2-} (Hopper and Parker 1999).

The ion interaction effects are required to interpret effects of soil properties on metal availability to plants. For example, increasing pH invariably decreases the free metal ion activity in soil (see previous section), which is in contrast to the above mentioned effects on the uptake of free metal ions from pore water into plants. It is tempting to predict the net effect of pH on metal bioavailability by properly ‘adding up’ the interactions on both sides. Such calculations are the basis of the *Biotic Ligand Model* (BLM – see next section) for predicting toxicity of metals and this model is an extension of the FIAM by taking ion interactions into account. The elegant study of Weng et al. (2003) is an example of this. That study showed that the nickel concentrations in oats grown at different pH were reasonably well described by combining solution culture data with data on free metal ion activities in pore water. The difference between predictions and observation were largest (factor 2) at lower nickel supply in soil, which may be due to lack of accounting for other ion interaction effects or lack of modelling rhizosphere conditions. A similar attempt was made to predict liming effects on cadmium uptake (Smolders and Helmke, unpublished). However, that study showed that the ion interactions studied separately did not add up, but predictions overestimated effects of liming (pH increase) and only correctly predicted the trend. The general trend emerging from large surveys are suggesting that the H^+ interaction for Cd^{2+} uptake is important. Field data on cadmium uptake by numerous plants (Table 8.4), for example, show that the net effect of increasing soil pH on reducing cadmium bioavailability is, on average, only a factor of 1.6 per unit pH increase. This impact is distinctly smaller than that effect on reducing solubility, i.e. a factor of 3.6 per unit on average (Degryse et al. 2009). Hough et al. (2005) similarly concluded from a large set of pot-trial data with ryegrass that H^+ decreased the availability of pore water Cd^{2+} and Zn^{2+} to an extent that the net effect of soil pH on decreasing crop cadmium is smaller than a factor of 1.5 per unit pH increase.

While the BLM concept improves our understanding of metal toxicity over the FIAM, it is still incomplete and not ready for practical use. First of all, the BLM is currently applied to pore water data while it is conceptually most correct to do that for the solution in the rhizosphere. The ionic composition in the rhizosphere

Table 8.4 The net statistical effect of soil pH on cadmium concentrations in vegetables based on a survey of crops and associated soils in contaminated and non-contaminated soils in Belgium and the Netherlands with pH roughly between 5–7. The effects given are all statistically significant at $P \leq 0.05$

Crop	n	Average factor decrease in crop Cd per unit pH increase at constant soil Cd
Potato tuber	239	1.2
Endive	76	2.1
Leek	139	1.9
Bean	48	2.1
Scorsonera	52	2.8
Lettuce	170	1.4
Spinach	96	1.1
Carrots	192	1.3
Maize	185	1.2
Celery	103	1.6
Grass	907	1.1
Overall		1.1–2.8, mean 1.6

Jansson et al. (2007)

differs dramatically from that in the bulk soil, in terms of concentrations of metals, accumulation of dissolved organic matter, accumulation of Ca^{2+} ions and pH gradients (McLaughlin et al. 1998). Secondly, nutrient solution culture has also its limitations as a model system for separately studying ion interactions: the ion activity ratios in nutrient solution are not necessarily identical to those at the root surface where the free metal ion activity is a complex function of uptake rate, diffusion and metal-ligand dissociation rate (see above). Thirdly, interactions in soil are usually complex with several factors involved. The effects of pH are interpreted and modelled as H^+:M^{2+} competition at the biological membrane, while the physiology behind this relationship is unclear. Modifying soil pH alters the availability of many elements that also can compete with the metal uptake process and it becomes tedious (and difficult) to take all such interactions into account. Finally, BLM modelling is not yet of practical use for Risk Assessment, both because it lacks precision (Antunes et al. 2006) and it requires numerous parameters (see below).

8.3.1.4 Translocation of Metals and Metalloids in the Plant

The long distance transport of solutes in plants takes place in the vascular system of xylem and phloem. Translocation is an important process in determining trace metal concentrations in plant tissues. For instance, Florijn and Van Beusichem (1993a) found that internal distribution rather than root uptake explained the genotypic differences in cadmium accumulation in shoots of maize inbreds. Figure 8.5 shows three typical patterns of metal partitioning between shoots and roots. In Fig. 8.5a, the root and shoot concentrations increase proportionally and are of similar magnitude

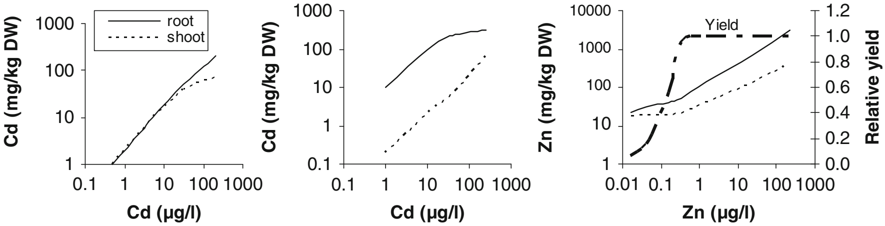


Fig. 8.5 Typical relationships between metal concentrations in the exposure solution and concentrations in shoot and root tissue for (**a**) a non-shoot cadmium excluder and (**b**) a shoot cadmium excluder (graphs based on results from Florijn and Van Beusichem, 1993b), and (**c**) for uptake of the essential element Zn. Panel (**c**) also shows the yield response (based on results of Degryse et al. 2006b)

('non-shoot excluder'; Florijn and Van Beusichem 1993a). In contrast, 'excluders' species retain most of the metals in the root, though the ability to retain the element in the root may strongly diminish beyond a threshold concentration (Fig. 8.5b). Figure 8.5c illustrates the partitioning of essential elements such as zinc. In the lower concentration range, the shoot concentrations are strongly regulated, and do not fall beneath a critical concentration. At low supply, the growth rate rather than the plant concentration decreases with decreasing metal supply.

Since the translocation to the shoot depends strongly on the genotype and external conditions – such as the availability of the metal – the translocation factor (usually defined as the ratio of shoot to root concentration) is not constant for a given element and plant species (Fig. 8.6). As a result, BCF values for vegetables expressed between shoot and soil reflect the combined impact of bioavailability of the contaminant in soil and the effect of soil conditions on the internal translocation. Table 8.5 gives a non-exhaustive overview of metal distribution in plants. This table illustrates the dependence of translocation on the metal, plant species and

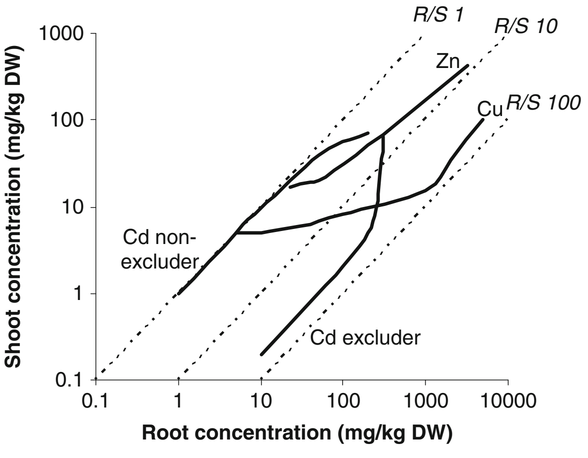


Fig. 8.6 Typical relationships between concentration in shoots and in roots for a non-shoot cadmium excluder, a shoot cadmium excluder, and for the essential elements zinc and copper (*full lines*) The *dotted lines* give the relations for the given ratios of root to shoot concentration (R/S)

Table 8.5 Concentrations of metal(loid)s in plant tissues for given concentrations of the element and other relevant conditions of the exposure solution. The ratio of the root to shoot concentration (R/S) is calculated based on concentration in the total shoot or in the leaves (when analyzed separately)

Element	Solution concentration (μg/L)	Other conditions	Plant species	Concentration (mg/kg_{dw}) in				R/S	References
				Roots	Stem or shoot	Leaves	Tubers		
Cd	1.1	Sand culture	Potato (Wilwash)	9	0.87	0.51	0.15	18	Dunbar et al. (2003)
			Potato (Kennebec)	2.7	0.63	0.28	0.24	10	
Cd	30	pH 4.5	Maize H96	75	0.5			150	Florijn and van Beusichem (1993b)
		pH 5.5		141	3.0			47	
		pH 6.5		210	5.9			36	
	30	pH 4.5	Maize B37	40	36			1.1	
		pH 5.5		60	55			1.1	
		pH 6.5		65	60			1.1	
Cd	2	pH 6	Ryegrass	255	5.7			45	Hatch et al. (1988)
			Cocksfoot	1021	40			26	
			Lettuce	48	14			3.4	
			Watercress	346	90			3.8	
Cd	14		Bean	96	2.4	2.0		49	Guo and Marschner (1995)
			Rice	262		2.6		100	
			Curly kale	114	2.8	3.4		34	
			Maize	114	43	47		2.4	
Ni	29		Bean	32	12	7.5		4.2	
			Rice	31		1.7		18	
			Curly kale	41	4.6	5.2		8.0	
			Maize	50	0.68	0.47		106	

Table 8.5 (continued)

Element	Solution concentration (μg/L)	Other conditions	Plant species	Concentration (mg/kg_{dw}) in Roots	Stem or shoot	Leaves	Tubers	R/S	References
Ni	252	pH 4	Oat	262	37			7.1	Weng et al. (2003)
		pH 5		523	59			8.9	
		pH 6		974	117			8.3	
		pH 7		2942	117			25	
Cd	1.2		Tomato	34	5.6			6.0	Checkai et al.f (1987)
	13			270	30			8.9	
Zn	7			183	98			1.9	
Cu	0.01			45	23			1.9	
Mo	9.6			13	14			0.9	
As(V)	749	10 μM P	Rice	1200	50			24	Geng et al. (2005)
		150 μM P		480	49			10	
As(V)	100	10 μM P	Velvetgrass	275	24			11	Quaghebeur & Rengel (2003)
		100 μM P		25	6			4	
Hg	100		Rice	300	2			150	Du et al. (2005)
	500			2600	18			144	
Hg	1204		Maize	1850	16			115	Rellan-Alvarez et al. (2006)
Hg	0^a	25% Hoagland	Water spinach	0.4	0.12	0.41		1.0	Göthberg et al. (2004)
	10.0			33	0.42	0.78		42	
	40.1			72	0.65	0.96		75	
Cd	0.02^a	25% Hoagland		0.13	0.05	0.04		3.3	
	101			676	59	30		23	
Pb	0^a	50% Hoagland		1.46	0.07	0.56		2.6	
	620^b			146	5.5	9.7		15	

Table 8.5 (continued)

Element	Solution concentration (μg/L)	Other conditions	Plant species	Concentration (mg/kg_{dw}) in Roots	Stem or shoot	Leaves	Tubers	R/S	References
Se(IV)	158	2 μM P	Ryegrass	632	197			3.0	Hopper and Parker (1999)
	158	2 μM P	Strawberry	537	55			9.7	
	395	2 μM P	Clover	632	70			9.0	
	395	200 μM P		434	79			5.5	
Se(VI)	395	0.5 mM SO_4^{2-}	Ryegrass	395	592			0.7	
	395	0.5 mM SO_4^{2-}	Strawberry	237	553			0.4	
	790	0.5 mM SO_4^{2-}	Clover	395	908			0.4	
	790	5 mM SO_4^{2-}		39	71			0.6	

[a] Not added to exposure solution; measured concentration below detection limit for Pb and Hg

[b] The added concentration was 5,000 μg/L, but the measured concentration was only 620 μg/L, presumably due to precipitation of lead phosphates

variety, external concentration of the element and solution composition (pH, other elements). Below we briefly discuss these major trends.

Metal or metalloid type, concentration and speciation: The results of Guo and Marschner (1995) clearly illustrate that translocation is element-specific. Nickel showed strong translocation to shoots in bean, curly kale, and to lesser extent, rice, and was excluded from the shoot of maize, while the opposite was observed for cadmium. The available data indicate that translocation of mercury to shoots is generally limited. The results of Göthberg et al. (2004) suggest that translocation of lead is relatively large at background concentrations. However, the much larger concentration in the leaves than in the stem suggests that lead in the leaves was mainly derived from aerial deposition. Most likely, translocation of lead from roots to shoots is limited, since the endodermis acts as a barrier (Sobotik et al. 1998), unless under specific conditions, e.g. after addition of synthetic chelators (EDTA) at high concentrations (Luo et al. 2005; Tandy et al. 2006b). Translocation of redox-sensitive metals or metalloids may depend on the redox state of the element in the pore water. Hopper and Parker (1999) found that translocation of selenium was much larger when the plants were exposed to selenate (Se(VI), SeO_4^{2-}) compared with selenite (Se(IV), $HSeO_3^-$) (Table 8.5). Wang et al. (2002) showed that the uptake rate of arsenic by *Pteris vittata* was smaller, but the translocation efficiency larger, for arsenite (As(III), H_3AsO_3) than for arsenate (As(V), $HAsO_4^{2-}/H_2AsO_4^-$).

Plant species and variety: Table 8.5 illustrates the large variation of cadmium translocation to shoots for plants that were grown under similar conditions (e.g. Guo and Marschner 1995; Hatch et al. 1988). Gramineae (maize, barley, oat, ryegrass, cocksfoot, rice) are often shoot cadmium excluders, though this cannot be generalized. For instance, some maize inbred lines show large cadmium translocation to shoot. Also for the Leguminosae (pea, beans, etc.), the translocation of cadmium to the aerial parts is usually small, whereas the Solanaceae (potato, tomato), Asteraceae (lettuce), Brassicaceae (cabbage, watercress) and Amaranthaceas (spinach) show in general relatively large translocation to the shoot (Kuboi et al. 1986).

Pore water composition: As discussed above, the availability of an element will affect its translocation. For instance, the translocation of essential metals (copper, zinc) is usually larger at small external concentration than at large external concentration (Fig. 8.6). The concentrations of other elements may also affect the translocation. For instance, Florijn and Van Beusichem (1993b) found that uptake of cadmium increased with increasing pH, presumably due to diminishing competition with H^+ ions for binding on the root surface. The translocation of cadmium to the shoot increased with increasing pH in the excluder plant species, whereas translocation was unaffected in the non-excluders (Table 8.5). Hatch et al. (1988) also found reduction of cadmium uptake with increasing pH, but found no consistent effect of pH on the translocation factor of cadmium in excluder (grasses) or non-excluder species. Increasing phosphate concentrations decreased the uptake of As(V), indicating competition for uptake between phosphate and arsenate, but increased the arsenic translocation to the shoot (Geng et al. 2005; Quaghebeur and Rengel 2003).

It is clear that many factors may affect the distribution of metals and metalloids within plants. Plants have a complex network of homeostatic mechanisms, including transport, chelation and sequestration, in order to maintain the concentrations of essential elements within the physiological limits and to minimize the detrimental effects of non-essential elements. The exact mechanisms are only starting to be unravelled. Most molecular insights obtained so far are on the cellular level, and very little is known about mechanisms controlling metal distribution on the level of the plant (Clemens 2001). The transport of metal(loid)s from roots to shoots is probably largely through the xylem. Guo and Marschner (1995) showed that the cadmium and nickel concentrations in the shoot dry matter were positively correlated with the concentrations in the xylem sap. The soluble fraction of cadmium in the roots was much larger for maize than for the other plant species tested, which was in agreement with its higher mobility in the plant (Table 8.5). Similarly, the higher soluble fraction of nickel in the roots extracts of bean compared with maize was in accordance with the higher nickel mobility in bean plants.

Phytochelatins (PCs) are metal-complexing peptides that play an important role in metal tolerance. The possible roles of PCs in heavy metal detoxification and homeostasis have been reviewed by Cobbett and Goldsbrough (2002). Phytochelatin synthesis is induced upon exposure to a variety of metals and metalloids (Grill et al. 1989). Overexpression of phytochelatin synthase was found to increase tolerance to cadmium, mercury and arsenate (Vatamaniuk et al. 1999). In some plants, sulphides also seem to play a role in the detoxification of cadmium by PCs. It was shown that phytochelatin-cadmium complexes of tomato grown at high cadmium concentration (100 μM) contained cadmium-S-peptide aggregates of ca. 2 nm diameter that consisted of a CdS crystallite core coated with PCs (Reese et al. 1992). Phytochelatins may be sequestered in the vacuole (Salt and Rauser 1995; Vogeli-Lange and Wagner 1990), but they may also be transported in the xylem (Gong et al. 2003). Translocation of cadmium in the xylem has been found to be independent of PC production (Florijn et al. 1993; Salt et al. 1995). Limited translocation of cadmium from shoot to root seems therefore not to be related to the presence of PCs, but is most likely due to sequestration of metals in the vacuoles of root cells, either as complexes with PC, as free ion or in another form. The transport from leaves to other plant tissues (e.g. grains, tubers) can occur in phloem only. For instance, cadmium in potato tubers and peanut kernels is not directly taken up from the soil, but is first transported in the xylem to the shoot, and then back down through the phloem (Popelka et al. 1996; Reid et al. 2003). Page and Feller (2005) showed that nickel and zinc were redistributed from older to younger leaves in wheat plants, indicating high phloem mobility, whereas Mn remained in the old leaves. Also split-root experiments and foliar application of ^{65}Zn demonstrated significant phloem transport of zinc from leaves to other plant parts (Haslett et al. 2001).

Differences in translocation of metals are an important factor determining the concentrations in plant tissues. The selection of plant species or cultivars with relatively small or large (in case of essential elements) concentrations in the harvested products can be used to manage metal concentrations in food crops. For instance, McLaughlin et al. (1994a) showed that potato cultivars grown commercially in Australia exhibited significant, nearly two-fold, differences in tuber cadmium concentration. Plant-breeding can also be an important tool to reduce the concentrations of potentially harmful elements, such as cadmium. For instance, low-cadmium durum wheat cultivars and sunflower hybrids have been developed for this purpose (Grant et al. 2008).

8.3.2 Foliar Uptake of Metals

In air, non-gaseous metals and metalloid are associated with small particulates (<10 μm) that can be deposited on plants. These *aerosols* may adsorb to plants or eventually also be absorbed and the metals and metalloids translocated. The mechanisms of the absorption processes are still unclear. Washing or even peeling vegetables in preparation for cooking does not fully remove the airborne metals (Dalenberg and Van Driel 1992). As a result, airborne metals can be a significant source of metals in the food chain. As shown below, this fraction can be even more important than that derived from soil.

Metals may also enter plants and the food chain through gaseous exposure through soil emissions (e.g. mercury). Little is known of the potential for uptake of metals by plants through gaseous routes of exposure (as opposed to aerosol or particulate exposure). Mercury (Hg) can be present in air as gaseous forms (Hg°) or as particulates (Hg-P). Plants can absorb mercury from soil *via* root uptake, from air as of Hg° *via* stomatal uptake and from air *via* adherence of Hg-P similar to the mechanisms described above. Plants can also be a source of mercury, releasing mercury when grown in low air mercury and high soil mercury. As a result, there is a so-called compensation point, the air concentration where no net flux of mercury vapour occurs and this point increases as soil mercury concentration increases (e.g. Ericksen and Gustin 2004).

There are very few data on uptake of airborne metals or metalloids and the methodologies to assess these are limited. Harrison and Chirgawi (1989) estimated the atmospheric contribution from the differences in metal (cadmium, chromium, nickel, lead and zinc) concentrations between plants grown in cabinets with either filtered or unfiltered air. An alternative method uses soils enriched with stable or radioactive isotopes and the isotope dilution in the crop as the basis for estimating airborne metal contribution (Mosbaek et al. 1989; Tjell et al. 1979). Dalenberg and Van Driel (1990, 1992) combined both methods to reduce the uncertainty related to the isotope ratio of bioavailable metals in soil. Finally, surveys of soil, plant and atmospheric concentrations combined with statistical tools allows a statistical, indirect, estimate of the fractions metal derived from atmosphere and soil (Voutsa et al. 1996). A compilation of data for cadmium and lead shows that most (generally >80%) of the lead in above-ground plant tissues of crops derives from the atmosphere (even in washed plant tissue), whereas variable results are obtained for cadmium (Table 8.6). The larger percentages for lead than for cadmium are related to the lower availability of soil lead relative to cadmium (largest sorption for lead). The airborne percentages of nickel and chromium are intermediate between cadmium and lead whereas the percentages for zinc are similar to those of cadmium (Harrison and Chirgawi 1989). The statistical method applied in an industrialized

Table 8.6 The fraction of airborne cadmium or lead in different crops calculated in experimental studies that used 'ambient air' (period 1975–1990) and 'background soils' Experiments were based on isotope labelling studies or comparison between crops grown in filtered air versus that in ambient air. All studies were performed < 1990, current air concentrations are lower and result in lower atmospheric contribution (see theoretical analysis in Table 8.7)

Crop	Percent of airborne	Air metal concentration ng m^{-3}	Soil metal concentration mg kg^{-1}	References
Cadmium				
Spinach	23	2.1	0.12–0.28	Harrison and Chirgawi (1989)
	n.s.	0.3–0.5	0.3	Dalenberg and van Driel (1990)
Lettuce	7–21	2.1	0.12–0.28	Harrison and Chirgawi (1989)
Wheat grain	21	0.3–0.5	0.3	Dalenberg and van Driel (1990)
	21	1.3	0.08	Hovmand et al. (1983)
Barley grain	41–58	1.3	0.08	Hovmand et al. (1983)
Rye grain	17–28	1.3	0.26	Hovmand et al. (1983)
Pea (peas)	0	2.1	0.12–0.28	Harrison and Chirgawi (1989)
Carrot root	4–8	2.1	0.12–0.28	Harrison and Chirgawi (1989)
	n.s.	0.3–0.5	0.3	Dalenberg and van Driel (1990)
	37–52	1.3	0.08	Hovmand et al. (1983)
Lead				
Grass	91–95	25–107	5	Dalenberg and van Driel (1990)
	90–99	132	11–17	Tjell et al. (1979)
Spinach	73	25–32	20	Dalenberg and van Driel (1990)
	85	110	11–41	Harrison and Chirgawi (1989)
	76–91	110	11–41	Harrison and Chirgawi (1989)

Table 8.6 (continued)

Crop	Percent of airborne	Air metal concentration ng m^{-3}	Soil metal concentration mg kg^{-1}	References
Wheat grain	100	25–32	20	Dalenberg and van Driel (1990)
Barley grain	93	40–90	11–17	Mosbaek et al. (1989)
Kale	97	40–90	11–17	Mosbaek et al. (1989)
Peas	0	110	11–41	Harrison and Chirgawi (1989)
Potato tuber	70	40–90	11–17	Mosbaek et al. (1989)
Carrot root	31	40–90	11–17	Mosbaek et al. (1989)
	32–44	110	11–41	Harrison and Chirgawi (1989)
	6–41	25–107	20	Dalenberg and van Driel (1990, 1992)

n.s. = not significant

area in Greece confirmed that concentrations of cadmium, chromium and lead in washed vegetables (cabbage, lettuce and endive) were mainly related to air contamination rather than soil contamination, and the same was found for arsenic (Voutsa et al. 1996). The consequence of the large atmospheric contribution is that crop metal concentrations can become unrelated to soil metal contamination, even when collating data with contrasting soil metal contaminations (Fig. 8.7). The relative contributions from soil and air obviously depend on the concentrations in air and soil, the bioavailability of soil metal or metalloid and the food preparation method (washing, boiling, etc.), i.e. generalizations cannot be made. A further consequence of foliar uptake is that soil remediation may not be completely successful for reducing metal concentrations in garden vegetables if atmospheric uptake is significant. Douay et al. (2008) examined the concentrations of metals in seven types of vegetables from remediated (complete soil replacement) and non-remediated soils in the vicinity of a metal smelter. Complete soil replacement reduced concentrations of cadmium in all vegetables (by 50–90%), but concentrations of lead were only reduced slightly, or not at all (Douay et al. 2008).

The transfer of metals and metalloids from air to plant can be aggregated in the so-called air accumulation factors (AAFs, with units m^3/g), calculated as the ratio of concentrations in the plant to that in air ($\mu g/g$ divided by $\mu g/m^3$). Other expressions relate the plant concentrations to metal deposition via bulk precipitation (g/ha/year).

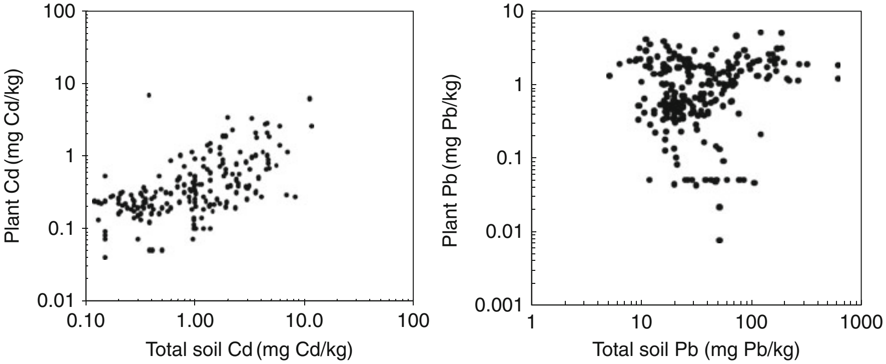


Fig. 8.7 Data of Dutch and Belgian surveys of metal concentration in field grown maize (whole shoots). Concentrations in maize generally correlate better with soil metal concentrations for cadmium (*left*) than for lead (*right*); lead concentrations in arial parts are often more reflecting air concentrations rather than soil concentrations, i.e. atmospheric contribution is larger than soil contribution in the case of cadmium. Data collated by Jansson et al. (2007) and Römkens and Rietra (unpublished)

Hovmand et al. (1983) used the isotope dilution method and bulk precipitation data to show that up to 70% of deposited cadmium could be incorporated in field-grown carrots, but the authors acknowledged that airborne submicron-sized particles containing cadmium, not sampled via the bulk precipitation, could also be filtered out by plant surfaces. Such interpretation suggests that AAF values may be a more suitable basis for modeling the transfer of metals and metalloids from air to plant. The AAF values observed by Harrison and Chirgawi (1989) ranged from 2 to 40 m^3/g for different plants and metals, typically about 10 m^3/g. Field-based estimates, using isotopically labeled soils, have yielded values that sometimes exceed 100 m^3/g for plant leaves (values for cadmium based on Hovmand et al. 1983). The AAF values are obviously not metal-specific values, but depend on plant type, aerosol properties, climatic factors, et cetera.

The above-mentioned data are at least 15 years old and there are little recent data measured under current atmospheric air quality, which typically demonstrates lower air metal concentrations in recent years due to environmental regulations (e.g. the ban on leaded gasoline). A simple sensitivity analysis allows modeling of the current atmospheric contribution, using the experimentally determined AAF and BCF values, the latter observed in 'clean air'. As shown in Table 8.7, we predict that still 50% of lead present in lettuce leaves in rural areas is derived from the atmosphere. A quantification of the pathway by which metals or metalloids enter the plant, i.e. by air or soil, has obvious consequences for Risk Assessment. For example, it is generally known that managing food chain contamination by metals could be achieved by managing air contamination for lead and, conversely, soil contamination for cadmium. However, the analysis in Table 8.7 suggest that human exposure to cadmium around smelters in the past, with cadmium concentrations in air often exceeding 100 ng/m^3, could have predominantly originated from atmospheric contributions to plants.

Table 8.7 A sensitivity analysis of predicted atmospheric contribution to plant metal concentrations, using current air quality conditions in rural and contaminated areas. Assumed air accumulation factors (AAF: plant:air concentration ratio, m^3/g_{dw}) and soil-plant bioconcentration factors (BCF, dry weight based) for lettuce are based on experimental data of Harrison and Chirgawi (1989). The 4 scenarios of soil and air concentrations are based on expert judgement, reflecting current air concentrations in rural areas in EU and background soil concentration for lowest values, and air concentrations near metal smelters and contaminated soils, for high concentrations

	AAF (m^3/g)	BCF (–)	Air conc. (ng/m^3)	Soil metal conc. mg/kg_{dw}	Predicted plant metal conc. (mg/kg_{dw})	Percent of airborne
Cd	20	1.5	0.2	0.2	0.3	1
	20	1.5	20	0.2	0.7	57
	20	1.5	0.2	2	3.0	0.1
	20	1.5	20	2	3.4	12
Pb	10	0.006	10	20	0.2	45
	10	0.006	200	20	2.1	94
	10	0.006	10	200	1.3	8
	10	0.006	200	200	3.2	62

8.4 Integrating Factors Affecting Metal/Metalloid Accumulation by Vegetables

There are a number of metal/metaloid-specific, soil and plant factors that can affect the accumulation of metals/metalloids by vegetables in urban gardens (Chaney 1973). Much of the information in this area is derived from studies of the use of sewage biosolids on land, with only a few studies examining contaminant uptake by vegetables in soils contaminated by industrial, mining or urban sources of contaminants. The most important factors controlling metal/metalloid accumulation by vegetables are outlined in the sections below.

8.4.1 Type of Metal/Metalloid

From the preceding sections, it is evident that the behaviour of metals/metalloids in urban/industrial soils and their uptake by plants are highly metal/metaloid specific. Cationic elements are generally more firmly held by soils than anions, and plants accumulate cations more readily than anions due to the negative electrical potential across the root membrane. Combining these properties of metals/metalloids, Chaney developed the 'soil-plant barrier concept' (Chaney 1980) where metals/metalloids were classified into groups for food-chain Risk Assessment. The classification of metals/metalloids depends on their partitioning behaviour in soil, their propensity for root uptake, and their propensity to accumulate in edible portions of plants in relation to critical phytotoxicity concentrations *versus* critical food concentrations for humans (Table 8.8). Using this concept, metals/metalloids which are highly insoluble, or are retained very strongly by plant roots, are in Group 1. Examples

Table 8.8 Maximum tolerable levels of metals/metalloids in plants in relation to maximum levels tolerated by animals in forages (modified from Chaney 1983)

	Level in plant foliage			Maximum levels chronically tolerated			
	Soil-plant barrier	Normal	Phytotoxic	Cattle	Sheep	Swine	Chicken
Metal(loid)	mg/kg$_{dw}$			mg/kg dry diet			
Group 1							
Cr^{3+}	yes	0.1–1	20	(3000)	(3000)	(3000)	3000
Group 2							
As, inorg.	yes	0.01–1.0	3–10	50	50	50	50
Pb	yes	2–5	–	30	30	30	30
Group 3							
B	yes	7–75	75	150	(150)	(150)	(150)
Cu	yes	3–20	25–40	100	25	250	300
Ni	yes	0.1–5	50–100	50	(50)	(100)	(300)
Zn	yes	15–150	500–1500	500	300	1000	1000
Group 4							
Cd	Fails	0.1–1	5–700	0.5	0.5	0.5	0.5
Co	Fail?	0.01–0.3	25–100	10	10	10	10
Mo	Fails	0.1–3.0	100	10	10	20	100
Se	Fails	0.1–2	100	(2)	(2)	2	2

of such metals are Cr^{3+} and also titanium, zirconium, tin, yttrium and silver. Group 2 comprises metals/metalloids that can be absorbed by roots, but are not readily translocated to edible plant parts, like arsenic, mercury and lead. Since 1980, it has become apparent that arsenic may be a food chain hazard in paddy rice production systems (due to the redox-induced mobilization of arsenic as described in Section 8.2.2), but in urban residential scenarios, arsenic can be regarded as a low hazard in terms of plant uptake. Group 3 comprises the elements boron, copper, nickel and zinc, which are easily taken up by plants, but are phytotoxic to plants before significant exposure to humans occurs. Group 4 are metals/metalloids which most likely pose a food-chain risk, like cadmium, cobalt, molybdenum and selenium, as plants can readily absorb and translocate these to edible portions, and they are not highly phytotoxic.

It should be noted that this approach classifies elements on the basis of transfer of contaminants via the soil-root-shoot/edible portion pathway. In urban contaminated soils, aerial deposition of contaminants (Section 8.4.2) provides a bypass of the ‘soil-plant barrier’, so that elements such as arsenic, lead and other metals/metalloids may pose a risk to human or animal health where there is a continuing source of metal/metalloid deposition to soil.

8.4.2 Vegetable Species

Vegetables vary widely in their ability to accumulate metals, either from soil or from atmospheric deposition. Numerous studies have examined this in both glasshouse trials, where soil-root uptake dominates, to field studies and surveys where both soil and atmospheric uptake pathways are important. Uptake by different vegetables varies according to metal/metalloid (and is not always consistent across sites), and in Table 8.9 vegetables are ranked according to their potential to accumulate contaminants. Cobalt and molybdenum are not listed, despite being in the group of elements likely to pose risks through food chain accumulation, due to the lack of data for these elements. It can be seen that in general leafy vegetables tend to be 'accumulators', and fruiting vegetables tend to be 'excluders' (Preer et al. 1980). This is likely because leaves accumulate ions through transpiration and aerial deposition, and non-essential elements may therefore accumulate in these tissues if the element is transported across the root membrane or deposited on the leaf surface. Fruits are principally formed by movement of nutrients in phloem (the nutrient transport system) into the fruit across a membrane from the xylem (the water transport system) within the plant, and transfer of many elements from xylem to phloem may be controlled, or affected by many counter ions, similar to root uptake processes (Section 8.4.1). For aerial deposition, which is mainly important for lead, species morphology (e.g. hairiness of leaves, surface roughness of aerial tissues) may be important in the trapping of atmospheric particulates and aerosols (Tiller et al. 1976, 1997).

8.4.3 Vegetable Cultivar

Most of the work on effects of vegetable cultivar on accumulation of metals or metalloids has focused on cadmium, since this is the element that is most likely to pose a food-chain hazard. There may be significant differences in cadmium uptake by different cultivars of the same vegetable species grown on a contaminated site, as was demonstrated for lettuce (Florijn et al. 1991; Thomas and Harrison 1989; Wang et al. 2007; Xue and Harrison 1991; Yuran and Harrison 1986), potato (Harris et al. 1981; McLaughlin et al. 1994b), curcurbits (Mattina et al. 2006), beans (Guo and Marschner 1996), carrots and peas (Alexander et al. 2006). Cultivar effects vary from insignificant up to a 3- to 4-fold difference in cadmium concentrations. Cultivar effects are therefore less pronounced than the influence of the type of vegetables, but selection of non-accumulating culivars may provide an important Risk Management action for minimizing exposure to humans through consumption of vegetables.

8.4.4 Soil Physical/Chemical Properties

In considering the effect of soil chemical and physical properties on metal/metalloid availability to plants, the type of contaminant, valence and charge, and other properties that determine fate in soil (Section 8.3) need to be considered. Anionic metalloids, for example, will respond differently in terms of crop accumulation compared to cationic metals. The principal factors governing vegetable uptake are discussed below.

Table 8.9 Relative differences in uptake of metals/metalloids by vegetable species. Note these may include both soil and atmospheric pathways. Studies focussing on contaminated urban soils were targeted, and where available, data from field-grown crops used

Element		As	Cd	Cr^{6+}	Hg	Pb	Se
Species	↑ Increasing accumulatation	Lettuce, radish, spinach	Lettuce, spinach	Spinach	Lettuce, spinach, cabbage	Lettuce, thyme	Garlic
		Onion, beetroot, carrot,	Cabbage, kale	Carrot	Onions, potato, carrot, garlic, collards	Turnip, cabbage	Onion, tomato, cabbage, lettuce, broccoli
		Peas, beans	Turnip	Beet	Tomato, beans, beet	Radish, carrot	Spinach
		Cauliflower, potato, tomato	Onion, leek, parsnip	Turnip, cabbage, snap bean, lettuce	Squash, pepper, cauliflower	Parsley, onion, spinach, beetroot, strawberry	Pepper, Squash, bean, beet, carrot
			Potato, tomato, zucchini, sweetcorn	Sweet corn, tomato		Corn, potato, tomato, raspberry, gooseberry, apple	
			French bean, beetroot, radish, pea, carrot, broccoli	Potato, cauliflower, peas, beans			
Sources		1	2	3	4	5	6

1 – Xu and Thornton (1985), Warren et al. (2003), Woolson (1973)
2 – Alloway et al. (1990), Preer et al. (1980), Davies and White (1981)
3 – Cary et al. (1977, 1996), Garcia et al. (1981)
4 – Cappon (1981, 1996)
5 – Dowdy and Larson (1975), Preer et al. (1980), Fleming and Parle (1977)
6 – Cappon (1981), Shane et al. (1988)

8.4.4.1 pH

Soil pH is perhaps the master soil variable governing the availability of metals and metalloids to plants (Page and Chang 1985). The major effect that pH has on both partitioning of metals/metalloids in soil (Section 8.3), and on response of plant roots to uptake of ions (Section 8.4.1), means that this variable accounts for most of the variation in metal-metalloid concentrations in vegetable tissues. Acidic soils will tend to increase uptake and accumulation of the cationic metals cadmium, copper, mercury, nickel, lead and zinc, while uptake and accumulation of the anionic elements arsenic, molybdenum and selenium will be reduced. As noted in Section 8.4.1, the effect of pH on accumulation of cationic metals is not predicted well just by consideration of changes in partitioning, as H^+ competition for root uptake mitigates the effects of pH on metal partitioning (Hough et al. 2005). Hence, liming of acidic soils will generally reduce uptake by vegetables of cationic metals in urban soils (Preer et al. 1980, 1995).

8.4.4.2 Soil Texture and Depth of Contamination

Soil texture (content of sand, silt and clay) will affect metal availability to crops as finer textured soils (clays) have a greater cation exchange capacity (CEC) and hence a greater ability to retain cationic metals (higher K_d) compared to sandy soils. Given the same total metal concentration in soil, clay-rich soils will produce crops with lower (cationic) metal concentrations. Depth of soil is also important as shallow contamination is likely to have a lesser effect on metal concentrations in vegetables compared to deeper contamination (Tiller et al. 1997), as plants may extend their root systems into less-contaminated subsoils.

8.4.4.3 Soil Organic Matter

Soil organic matter (OM) plays a similar role to clays in affecting metal concentrations in vegetables, as OM is a major contributor to the pH-dependent negative charge in soils which gives rise to soils' ability to retain cationic metals (represented by the CEC). Addition of OM to soil in the form of compost can therefore markedly reduce cationic metal uptake by plants (Farfel et al. 2005; Jones et al. 1987; Narwal and Singh 1998; Traulsen and Schonhard 1987; Verloo and Willaert 1988). Addition of OM has also been found to reduce arsenic accumulation by vegetables from contaminated soils (Cao and Ma 2004).

Many vegetable gardens use imported OM (e.g. composts, mulches or other soil amendments) and these may be useful in minimising uptake of cationic metals by vegetable crops. Also, high OM contents in many garden soils compared to agricultural soils means that soil to plant transfer factors developed from agricultural surveys (which may not identify OM as an important factor controlling plant uptake) may overestimate metal uptake in garden scenarios.

However, it should also be noted that high concentrations of dissolved OM in a soil may lead to increased potential for leaching of cationic metals, due to complexation by dissolved OM reducing K_d values (Sauvé et al. 2000). This is particularly important for copper.

8.4.4.4 Salinity

Salinity may play a role in enhancing uptake of cadmium from soil by vegetables. Chloro-complexation of Cd^{2+} ions reduces the charge of the cadmium ion ($CdCl^+$ or $CdCl_2{}^0$) and hence increases its mobility through soil (Hahne and Kroontje 1973). Chloro-complexation also increases the diffusive flux of cadmium to root surfaces (Degryse et al. 2006a; Smolders and McLaughlin 1996a, 1996b). As a result of this increased mobility in soil and increased diffusive flux of cadmium to root uptake sites, cadmium concentrations in vegetables are increased when soil salinity increases, and this has been demonstrated in field-grown crops (Helal et al. 1998; McLaughlin et al. 1994a, 1997).

Theoretically, soil salinity should also have a large effect on uptake of mercury, but this has not been demonstrated in any uptake studies by vegetables in soil. Effects of salinity on the uptake of other metals should be minimal, and for the anionic metals/metalloids, the chloride ion may actually inhibit uptake of these ions, although this has not been demonstrated in practice.

8.4.4.5 Redox Potential

Redox potential, through the effects on reductive dissolution of Fe and Mn oxides (Section 8.2.2), can have a marked influence on the uptake of several metals by plants. For example, uptake of arsenic by rice and uptake of cobalt by pasture plants (Adams and Honeysett 1964) has been found to be significantly enhanced under low redox conditions, again due to reductive dissolution of Fe/Mn oxides, and in the case of cobalt, reductive dissolution of insoluble Co^{3+} to soluble Co^{2+}.

There is little information in the literature on effects of low redox on metal uptake by vegetables. It should be noted that most plants are not tolerant of waterlogged soil conditions which create low redox potentials, and generally die due to lack of oxygen. Redox may therefore not be a major risk factor for affecting metal uptake by vegetables in urban gardens.

8.4.4.6 Nutrient Status

It is well known that soil nutrient status can influence metal uptake by plants, and the effects can operate in several ways. Low levels of addition of nitrogen and phosphorus in soil have been shown to enhance cadmium uptake by crops (Grant and Sheppard 2008; Maier et al. 2002a, b; Williams and David 1973). Higher rates of addition of N fertiliser are likely to acidify soils and lead to enhanced uptake of cationic metals (Eriksson 1990; Grant et al. 1996; Lorenz et al. 1994), while high rates of P addition may reduce uptake of cationic metals through precipitation reactions (McGowen et al. 2001). High P addition, however, may promote uptake of anionic metals/metalloids, due to displacement of these ions from sorbing surfaces in soil (Cao and Ma 2004; Creger and Peryea 1994). High additions of potassium (K) fertiliser as KCl can enhance cadmium uptake by crops due to the chloro-complexation issue described in Section 8.4.4.4 (Sparrow et al. 1994).

In many urban garden soils, fertilisers are often added in excess of plant requirements, so this factor may play an important role in metal accumulation by vegetables. In particular, it is important that N (from either manufactured fertilisers or organic sources of N) is not used in excess, as this rapidly acidifies the soil and will enhance cationic metal uptake by vegetables.

8.5 Models to Predict Contaminant Uptake by, or Toxicity to, Vegetables

Different models with increasing complexity can be used to predict metal concentrations in vegetables (Table 8.10). Among the easiest to apply are those that assume a constant plant metal concentration or the ones that relate the plant metal concentration to certain soil parameters: an equilibrium approach. In more process-oriented approaches, the concentration is not necessarily constant throughout the growing period. Here also the uptake of metals is a function of certain soil parameters, but also specific plant parameters that for instance regulate water uptake, such as root growth and length.

An important aspect related to Risk Assessment is that only the metal concentrations in the edible products of the plant (root, stem or fruit) are considered relevant, although metals are present in all plant parts. Identification of those processes responsible for uptake of metals have been described in papers such as by Clemens (2001) and Hall (2002) and will not be dealt with here. Also models that

Table 8.10 Overview of model approaches to predict heavy metal contents in vegetables

Type of approach	Examples	Scale of use	Advantages/problems
1. Constant heavy metal content		Site-specific Risk Assessment	
2. Soil-plant transfer models	BAF, BCF, regression models	Site-specific/national Risk Assessment	Advantage: available soil parameters
3. Free ion activity model	FIAM		
4. Biotic ligand model	BLM	Experimental	Problem: parameters are not available at regional/national scale,
5. Physiological models		Site-specific	Advantage: predicts differences between years
6. Barber-Cushman mechanistic model		Experimental	

are used to predict toxicity for plants (e.g., Thakali et al. 2006) will not be dealt with here. Although models to predict toxicity are similar to metal uptake models, they do not describe the metal concentrations or translocation to the food product.

The metal uptake models differ strongly and vary in complexity from very simple models to detailed growth models with a high requirement for input data. Also the scale from which they have been used differs from pot experiments to models used on a national scale (Table 8.10).

8.5.1 Model Characteristics

8.5.1.1 Constant Heavy Metal Content for Each Plant Species

This assumption is used in most site-specific studies when the metal concentration is assessed by actual site measurements (e.g. Sipter et al. 2008). In principle it is assumed that the metal concentrations do not differ within a year, between years and are not influenced by the applications of soil amendments e.g. lime or changes in vegetable type over time.

8.5.1.2 Soil-Plant Transfer Models

A simple and commonly-used approach is the bioconcentration factor (BCF) or bioaccumulation factor (BAF). The terms BCF and BAF originate from studies of contaminant concentrations in water (mg/L) and biota (mg/kg$_{\mathrm{dw}}$). Originally BCF or BAF are expressed as the ratio between the concentration in the biota and water and are therefore given in L/kg. However several soil Risk Assessment models such as CLEA (Environmental Agency 2008) and CSOIL (Brand et al. 2007) use dimensionless BCF or BAFs, most commonly expressed as the ratio between the metal/metalloid concentration in the vegetable (mg/kg) and that in soil (mg/kg). Compilations of BCF values are given by Sauerbeck and Styperek (1985) and Bechtel Jacobs (1998). The metal concentration is either expressed on a fresh or dry weight basis. The BCF approach is rather easy to apply, but it has been shown that BCF values are not constant, but depend on the level of soil contamination (see Section 8.4.1.1). Often BCF values decrease with increasing levels of soil contamination (Wang et al. 2004).

Metal concentrations in plant tissues are often linked to those in various soil extracts by regression analysis. Linear or log transformed data for metal concentrations in soil, soil solution or soil extracts are linked to measured plant concentrations. For metals like cadmium in the soil (M_{soil}) and the plant (M_{plant}), Freundlich type functions are often used (Efroymson et al. 2001; Krauss et al. 2002).

$$M_{\mathrm{plant}} = 10^{\mathrm{a}} M_{\mathrm{soil}}^{\mathrm{b}} \text{ or } \log\left[M_{\mathrm{plant}}\right] = \mathrm{a} + \mathrm{b}\log[M_{\mathrm{soil}}] \tag{8.2}$$

where M_{plant} is metal concentration in the plant, M_{soil} is metal concentration in the soil, and a and b are constants.

The Freundlich-type equation can be extended, using soil parameters such as soil pH and soil organic matter (S) (Adams et al. 2004; Efroymson et al. 2001; Hough et al. 2003; Plette et al. 1999).

$$\log [M_{\text{plant}}] = \text{a} + \text{b} \log [M_{\text{soil}}] + \text{c} [pH] + \text{d} \log [S] \tag{8.3}$$

Efroymson et al. (2001) have shown that incorporation of soil pH improved the model for cadmium, mercury, selenium and zinc. Incorporation of other factors such as CEC, total metal concentrations in soil and extractable concentrations in soil using Freundlich-type equations has been reported for various vegetables and metals (Wang et al. 2004). In most cases the use of soil parameters, like pH, organic matter, CEC or texture, improves the model performance compared to those based on extractable or total concentrations only. For example in Fig. 8.8 and in Table 8.11

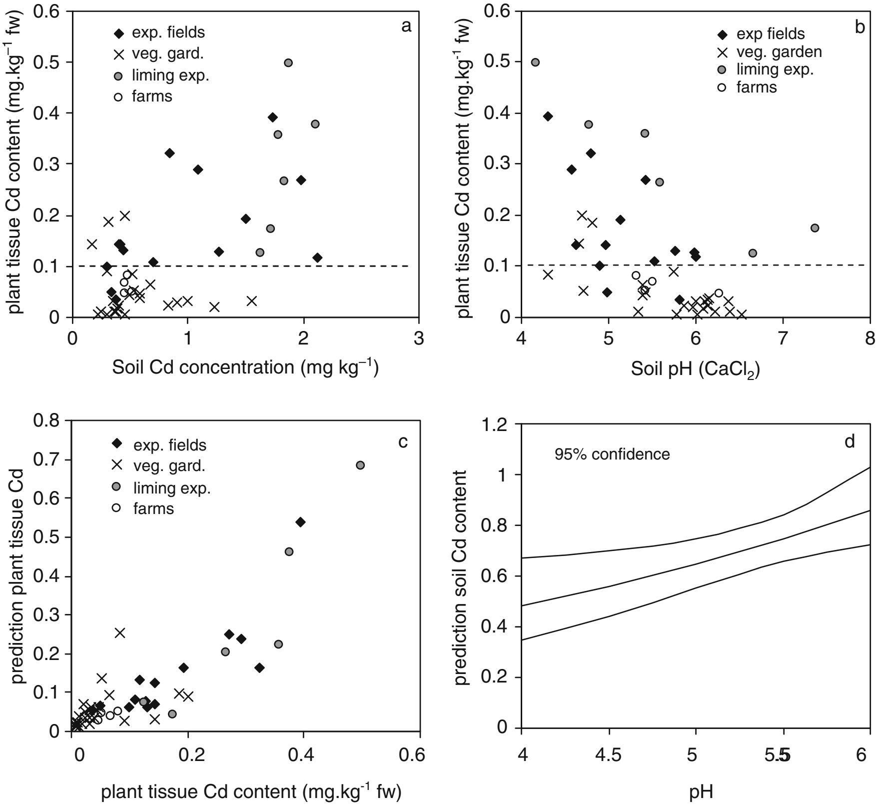


Fig. 8.8 Concentrations of cadmium in leek (*Allium ampeloprasum* L.) as a function of soil cadmium concentration (0.43 M HNO_3 extraction) (**a**), as a function of soil pH (**b**), predicted versus measurements (**c**), and predicted soil concentrations (model No. 4 in Table 8.11) as a function of pH where the limit value for cadmium in leek is exceeded (**d**). Results are from several studies near the Dutch zinc smelter in Budel: experimental fields, vegetable gardens, liming experiments and farms. The dotted line is the European limit value for cadmium in leek

Table 8.11 Soil to plant transfer models for cadmium in leek on the basis of Eq. (8.3) for the data in Fig. 8.8. Given for each model are the coefficients of determination (R^2), and standard errors (se). SOM = soil organic matter content (%)

Nr.	Regression models	R^2	se
1	$\text{Log } M_{\text{plant}} = -1 + 0.89 \log M_{\text{soil}}$	0.29	0.43
2	$\text{Log } M_{\text{plant}} = 1.0 - 0.39\ \text{pH}$	0.30	0.42
3	$\text{Log } M_{\text{plant}} = 1.32 + 0.95 \log M_{\text{soil}} - 0.42\ \text{pH}$	0.63	0.31
4	$\text{Log } M_{\text{plant}} = 1.64 + 1.21 \log M_{\text{soil}} - 0.37\ \text{pH}$ $-0.87 \log \text{SOM}$	0.67	0.30

it can be seen that the cadmium concentrations in leek in the sandy soils in the Belgian-Dutch Kempen region depend on the cadmium concentration of the soil and soil pH. Using Eq. 8.3, with soil cadmium, soil pH and soil organic matter, most of the variation can be explained. Such a model can be used, for example, to predict the threshold soil cadmium concentrations as a function of pH above which the food cadmium standards would be exceeded.

An advantage of Freudlich-type soil-plant transfer relations is the simplicity and the applicability. Most equations use variables that are available from soil investigations, such as total metal content, pH, organic matter and CEC. However these soil-plant equations should not be used for soils where concentrations of metals are outside the range from which the regressions were derived. Römkens et al. (2009) studied the quality of the soil-plant transfer equations for rice and concluded that only models for cadmium and zinc gave good predictions. While predictions for cobalt and nickel where not as good, and prediction for copper and lead were not possible.

8.5.1.3 FIAM

The free ion activity model (FIAM) (Morel 1983) is based on the assumption that metal uptake or toxicity is related directly to the free ion activity in the pore water. While early reports suggested this was a promising model (Sauvé et al. 1996), later studies have questioned whether free ion activities alone can improve predictions of plant metal uptake, given effects of diffusional limitations adjacent to and near the plant root (Degryse et al. 2006a, b; Hudson 2005; McLaughlin 2001a) and ion competition and other factors that can markedly affect plant metal/metalloid acquisition (Section 8.4.1). Free metal activities could not explain nickel uptake by oats in glasshouse experiments (Weng et al. 2003, 2004) nor copper uptake and toxicity to barley and tomato, also in glasshouse experiments (Zhao et al. 2004). Nolan et al. (2005) examined a wide range of techniques to predict cationic metal (cadmium, copper, lead, and zinc) uptake by wheat (again in glasshouse trials), including free metal ion activities, and found consideration of elemental free ion activities did not improve the predictions. McLaughlin et al. (1997) examined the relationship between free Cd^{2+} ion activities and concentrations of cadmium in potato tubers grown in field soils, and again found that consideration of free ion activities did not improve prediction of metal accumulation in tubers. This is not to say that knowledge of elemental speciation is not important in explaining metal accumulation by plants, but the analytical effort of separating and determining free metal ion activities in pore water may not be justified in terms of improving predictions of metal uptake by plants. This is more fully explained in Nolan et al. (2003).

8.5.1.4 Biotic Ligand Model

Originally the BLM described metal uptake and toxicity as a function of binding to certain biotic ligands present on surfaces of aquatic organisms (Di Toro et al. 2001). The applicability of the BLM has been studied for a range of aquatic organisms and heavy metals (lead, copper, zinc, nickel) and has shown potential to be used to define water quality standards (Comber et al. 2008). Toxicity has been described for soils using a BLM approach for copper and nickel toxicity to barley (Thakali et al. 2006), or nickel toxicity to oat (Weng et al. 2003; 2004). However zinc uptake by algae could not be predicted by surface-bound zinc or solution zinc chemistry, probably due to internal regulation of zinc (Hassler and Wilkinson 2003). Also the competition between different cations in a BLM could not improve the prediction of cadmium and zinc uptake by *Lolium perenne* (Hough et al. 2005), or describe the competition between copper and lead and zinc on algae (Hassler et al. 2004).

A problem of the FIAM and the BLM is that if these models are used to predict effects of various cations on metal uptake by roots (Cheng and Allen 2001), one still needs to describe the translocation of the metal from root to shoot or fruit with a translocation coefficient, as was done by Cheng and Allen (2001) for copper. However, translocation of metals from root to shoot is not constant, especially not for essential metals such as copper and zinc (Kalis et al. 2006).

8.5.1.5 Physiological Models

Uptake of any contaminants from soil is related to the water uptake and, for some contaminants also the concentration or free ion activity in the pore water (Ingwersen and Streck 2006; Peijnenburg et al. 2000). In models for neutral organic xenobiotics it is often assumed that uptake is a passive process related to water transpiration (see Chapter 9 by Trapp and Legind, this book). Ultimately, the metal content of plant tissues is a function of plant growth, water transpiration, metal concentration or speciation in the pore water, selectivity/ion competition effects at the root surface and diffusional limitations to uptake. An advantage of such models is that they can explain differences in metal uptake between years based on differences in weather. In practice this still seems to be rather complex, because other parameters (e.g. root depth) also vary between years (Ingwersen and Streck 2006). The physiological model can be extended to include competitive effects of other cations, such as the effect of H^+ and Zn^{2+} on Cd^{2+} uptake by plants, similar to BLMs (Ingwersen and Streck 2006).

8.5.1.6 Barber-Cushman Mechanistic Model

A commonly-used model to predict the uptake of nutrients by plants is the *Barber-Cushman model* (Barber 1995). This model has also been used to describe the uptake of heavy metals, including uptake of zinc by rice (Adhikari and Rattan 2000), cadmium by maize and *Thlaspi* (Sterckeman et al. 2004) and zinc by *Thlaspi* (Whiting et al. 2003). The Barber-Cushman model includes plant and soil parameters, such as root geometry, growth and kinetic parameters for the uptake of ions, parameters which can only be determined in detailed, small scale experiments. In general, application of this kind of model helps to identify which factors determine metal uptake by plants from a research perspective, but is less useful for generic predictive modelling.

8.5.2 Application of Models

In most Risk Assessment studies, soil-plant transfer models have been used. Also the physiological model of Ingwersen and Streck (2006) has been applied in a small region. Human exposure due to vegetable consumption has been described in Elert et al. (Chapter 11 of this book). Soil-plant transfer models have been used in:

- large-scale Risk Assessment on the basis of national soil data (Brus and Jansen 2004; Brus et al. 2002; De Vries et al. 2008);
- derivation of critical soil concentrations (Brus et al. 2005; De Vries et al. 2007) on the basis of limit values for crops and fodder;
- local and regional Risk Assessment on the basis of standard soil analysis in CLEA or CSOIL (Römkens et al. 2005); and
- local and regional Risk Assessment on the basis of soil analysis and soil maps (Hough et al. 2004).

Soil-to-plant transfer models for metals can be applied in Risk Assessment studies on a regional and local scale, either as such or in combination with additional sampling of soil and vegetables. However, parameters of current models very much depend on soil and plant data on which they are based which renders most models suitable for local applications only. The applicability of generic soil-to-plant transfer models to predict heavy metal contents in local circumstances should be checked. Ideally the models are verified with measurements on soil and plants. A practical use of the models in Risk Assessment of vegetable gardens is, for example, to calibrate the models for the crops which have the highest contribution to metal intake by humans (e.g. potato, leafy vegetables) and to use generic models for all other crops. In the next chapters it will be shown that the use of the model also depends on the questions asked.

References

Adams SN, Honeysett JL (1964) Some effects of soil waterlogging on the cobalt and copper status of pasture plants grown in pots. Austr J Agric Res 15:357–367

Alexander PD, Alloway BJ, Dourado AM (2006) Genotypic variations in the accumulation of Cd, Cu, Pb and Zn exhibited by six commonly grown vegetables. Environ Pollut 144:736–745

Alloway BJ, Jackson AP, Morgan H (1990) The accumulation of cadmium by vegetables grown on soils contaminated from a variety of sources. Sci Total Environ 91:223–236

Anderson PR, Christensen TH (1988) Distribution coefficients of Cd, Co, Ni, and Zn in soils. J Soil Sci 39:15–22

Antunes PMC Hale BA (2006) The effect of metal diffusion and supply limitations on conditional stability constants determined for durum wheat roots. Plant Soil 284:229–241

Antunes PMC, Berkelaar EJ, Boyle D, Hale BA, Hendershot W, Voigt A (2006) The biotic ligand model for plants and metals: technical challenges for field application. Environ Toxicol Chem 25(3):875–882

Adams ML, Zhao FJ, McGrath SP, Nicholson FA Chambers BJ (2004) Predicting cadmium concentrations in wheat and barley grain using soil properties. J Environ Qual 33:532–541

Adhikari T, Rattan RK (2000) Modelling zinc uptake by rice crop using a Barber-Cushman approach. Plant Soil 227:235–242

Barber SA (1995) Soil nutrient bioavailability. A mechanistic approach. Wiley, New-York, 414 pp

Bartlett RJ, James BR (1988) Mobility and bioavailability of chromium in soils. In: Nriagu JO, Nieboer E (eds) Chromium in the natural and human environments. Wiley Interscience, New York, pp 267–304

Bechtel Jacobs Company LLC (1998) Empirical models for the uptake of inorganic chemicals from soil by plants. Bechtel Jacobs Company LLC, Oak Ridge. TN. BJC/OR-133

Bowell RJ (1994) Sorption of arsenic by iron-oxides and oxyhydroxides in soils. Appl Geochem 9:279–286

Brand E, Otte PF, Lijzen JPA (2007) CSOIL 2000: an exposure model for human risk assessment of soil contamination: A model description RIVM report 711701054, Bilthoven, The Netherlands

Buchter B, Davidoff B, Amacher MC, Hinz C, Iskandar IK, Selim HM (1989) Correlation of Freundlich *Kd* and *n* retention parameters with soils and elements. Soil Sci 148:370–379

Brown SL, Chaney RL, Angle JS, Ryan JA (1998) The phytoavailability of cadmium to lettuce in long-term biosolids-amended soils. J Environ Qual 27:1071–1078

Brus DJ, Jansen MJW (2004) Uncertainty and sensitivity analysis of spatial predictions of heavy metals in wheat. J Environ Qual 33(3):882–890

Brus DJ, de Gruijter JJ, Romkens PFAM (2005) Probabilistic quality standards for heavy metals in soil derived from quality standards in crops. Geoderma 128(3–4):301–311

Brus DJ, de Gruijter JJ, Walvoort DJJ, de Vries F, Bronswijk JJB, Romkens PFAM, de Vries W (2002) Mapping the probability of exceeding critical thresholds for cadmium concentrations in soils in the Netherlands. J Environ Qual 31(6):1875–1884

Cao X, Ma LQ (2004) Effects of compost and phosphate on plant arsenic accumulation from soils near pressure-treated wood. Environ Pollut 132:435–442

Cappon CJ (1981) Mercury and selenium content and chemical form in vegetable crops grown on sludge-amended soil. Archiv Environ Contamin Toxicol 10:673–689

Cary EE, Allaway WH, Olson OE (1977) Control of chromium concentrations in food plants.1. Absorption and translocation of chromium by plants. J Agric Food Chem 25:300–304

Chaney RL (1973) Crop and food chain effects of toxic elements in sludges and effluents. Proceedings of the joint conference on "recycling municipal sludges and effluents on land." National Association of State Universities and Land-Grant Colleges, Washington, DC, pp 129–141

Chaney RL (1980) Health risks associated with toxic metals in municipal sludge. In: Bitton G, Damro BL, Davidson GT, Davidson JME (eds) Sludge – health risks of land application. Ann Arbor Sci, MI, USA, pp 59–83

Chaney RL (1983) Potential effects of waste constituents on the food chain. In: Parr JF, Marsh PB, Kla JM (eds) Land treatment of hazardous waste Noyes Data Corp. Park Ridge, NJ, pp 152–240

Creger TL, Peryea FJ (1994) Phosphate fertilizer enhances arsenic uptake by apricot liners grown in lead-arsenate-enriched soil. HortScience 29:88–92

Checkai RT, Corey RB, Helmke PA (1987) Effects of ionic and complexed metal concentrations on plant uptake of cadmium and micronutrient metals from solution. Plant Soil 99:335–345

Clemens S (2001) Molecular mechanisms of plant metal tolerance and homeostasis. Planta 212(4):475–486

Chen WP, Li LQ, Chang AC, Wu LS, Chaney RL, Smith R and Aiwa H (2009) Characterizing the solid-solution partitioning coefficient and plant uptake factor of As, Cd, and Pb in California croplands. Agric Ecosyst Environ 129:212–220

Cheng T, Allen HE (2001) Prediction of uptake of copper from solution by lettuce (*Lactuca sativa* Romance). Environ Toxicol Chem 20:2544–2551

Cobbett C, Goldsbrough P (2002) Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annu Rev Plant Biol 53:159–182

Collins RN, Onisko BC, McLaughlin MJ, Merrington G (2001) Determination of metal-EDTA complexes in soil solution and plant xylem by ion chromatography-electrospray mass spectrometry. Environ Sci Technol 35:2589–2593

Comber SDW, Merrington G, Sturdy L, Delbeke K, van Assche F (2008) Copper and zinc water quality standards under the EU Water Framework Directive: the use of a tiered approach to estimate the levels of failure. Sci Total Environ 403:12–22

Dalenberg JW, van Driel W (1990) Contribution of atmospheric deposition to heavy-metal concentrations in field crops. Netherlands J Agric Sci 38:369–379

Dalenberg JW, van Driel W (1992) De directe bijdrage van Pb en Cd in atmosferische depositie aan de Pb-en Cd-belasting van gras, spinazie, waspeen en tarwe in een zandige zavelgrond. DLO-Instituut voor Bodemvruchtbaarheid, Haren, the Netherlands

Davies BE, White HM (1981) Trace elements in vegetables grown on soils contaminated by base metal mining. J Plant Nutr 3:387–396

Degryse F, Smolders E, Merckx R (2006a) Labile Cd complexes increase Cd availability to plants. Environ Sci Technol 40:830–836

Degryse F, Smolders E, Parker DR (2006b) Metal complexes increase uptake of Zn and Cu by plants: implications for uptake and deficiency studies in chelator-buffered solutions. Plant Soil 289:171–185

Degryse F, Smolders E, Parker DR (2009) Partitioning of trace metals in soils: concepts, methodologies, predictions and applications – a review. Eur J Soil Sci 60:590–612

Degryse F, Verma VK, Smolders E (2008) Mobilization of Cu and Zn by root exudates of dicotyledonous plants in resin-buffered solutions and in soil. Plant Soil 306:69–84

Di Toro DM, Allen HE, Bergman HL, Meyer JS, Paquin PR, Santore RC (2001) Biotic ligand model of the acute toxicity of metals. I. Technical Basis. Environ Toxicol Chem 20: 2383–2396

Douay F, Roussel H, Pruvot C, Loriette A, Fourrier H (2008) Assessment of a remediation technique using the replacement of contaminated soils in kitchen gardens nearby a former lead smelter in Northern France. Sci Total Environ 401:29–38

Dowdy RH, Larson WE (1975) The availability of sludge-borne metals to various vegetable crops. J Environ Qual 4:278–282

Du X, Zhu YG, Liu WJ, Zhao XS (2005) Uptake of mercury (Hg) by seedlings of rice (Oryza sativa L.) grown in solution culture and interactions with arsenate uptake. Environ Exp Bot 54: 1–7

Dunbar KR, McLaughlin MJ, Reid RJ (2003) The uptake and partitioning of cadmium in two cultivars of potato (*Solanum tuberosum* L.). J Exp Bot 54:349–354

Efroymson RA, Sample BE, Suter GW II (2001) Bioaccumulation of inorganic chemicals from soil by plants: regressions of field data. Environ Toxicol Chem 20:2561–2571

Eriksson JE (1990) Effects of nitrogen-containing fertilizers on solubility and plant uptake of cadmium. Water Air Soil Pollut 49:355–368

Ericksen JA, Gustin MS (2004) Foliar exchange of mercury as a function of soil and air mercury concentrations. Sci Total Environ 324:271–279

Farfel MR, Orlova AO, Chaney RL, Lees PSJ, Rohde C, Ashley PJ (2005) Biosolids compost amendment for reducing soil lead hazards: a pilot study of Orgro amendment and grass seeding in urban yards. Sci Total Environ 340:81–95

Fleming GA, Parle PJ (1977) Heavy-metals in soils, herbage and vegetables from an industrialized area west of Dublin-city. Irish J Agric Res 16:35–48

Florijn PJ, Nelemans JA, Van Beusichem ML (1991) Cadmium uptake by lettuce varieties. Netherlands J Agric Sci 39:103–114

Florijn PJ, Van Beusichem ML (1993b) Cadmium distribution in maize inbred lines: effects of pH and level of Cd supply. Plant Soil 153:79–84

Florijn PJ, van Beusichem ML (1993a) Uptake and distribution of cadmium in maize inbred lines. Plant Soil 150:25–32

Florijn PJ, De Knecht JA, van Beusichem ML (1993) Phytochelatin concentrations and binding state of Cd in roots of maize genotypes differing in shoot root Cd partitioning. J Plant Physiol 142:537–542

Garcia WJ, Blessin CW, Inglett GE, Kwolek WF, Carlisle JN, Hughes LN, Meister JF (1981) Metal accumulation and crop yield for a variety of edible crops grown in diverse soil media amended with sewage sludge. Environ Sci Technol 15:793–804

Geng CN, Zhu YG, Liu WJ, Smith SE (2005) Arsenate uptake and translocation in seedlings of two genotypes of rice is affected by external phosphate concentrations. Aquat Bot 83:321–331

Goldberg S (1997) Reactions of boron with soils. Kluwer Academic, Netherlands, pp 35–48

Goldberg S, Forster HS (1998) Factors affecting molybdenum adsorption by soils and minerals. Soil Sci 163:109–114

Gong JM, Lee DA, Schroeder JI (2003) Long-distance root-to-shoot transport of phytochelatins and cadmium in Arabidopsis. Proc Natl Acad Sci USA 100:10118–10123

Grant CA, Bailey LD, Therrien MC (1996) The effect of N, P and KCl fertilizers on grain yield and Cd concentration of malting barley. Fertilizer Res 45:153–161

Göthberg A, Greger M, Holm K, Bengtsson BE (2004) Influence of nutrient levels on uptake and effects of mercury, cadmium, and lead in water spinach. J Environ Qual 33:1247–1255

Grant CA, Sheppard SC (2008) Fertilizer impacts on cadmium availability in agricultural soils and crops. Human Ecol Risk Assess 14:210–228

Grant CA, Clarke JM, Duguid S, Chaney RL (2008) Selection and breeding of plant cultivars to minimize cadmium accumulation. Sci Total Environ 390:301–310

Grill E, Loffler S, Winnacker EL, Zenk MH (1989) Phytochelatins, the heavy-metal-binding peptides of plants, are synthesized from glutathione by a specific gamma-glutamylcysteine dipeptidyl transpeptidase (Phytochelatin Synthase). Proc Natl Acad Sci USA 86:6838–6842

Guo YT, Marschner H (1995) Uptake, distribution, and binding of cadmium and nickel in different plant species. J Plant Nutr 18:2691–2706

Guo Y, Marschner H (1996) Genotypic differences in uptake and translocation of cadmium in bean and maize inbred lines. Zeitschrift fuer Planzenernaehrung und Bodenkunde 159:55–60

Haghiri F (1973) Cadmium uptake by plants. J Environ Qual 2:93–95

Hahne HCH, Kroontje W (1973) Significance of pH and chloride concentration on behaviour of heavy metal pollutants: mercury (II), cadmium (II), zinc (II), and lead (II). J Environ Qual 2:444–450

Hall JL (2002) Molecular mechanisms of plant metal tolerance and homeostasis. J Exp Bot 53 (366):1–11

Harris MR, Harrison SJ, Wilson NJ, Lepp NW (1981) Varietal differences in trace metal partitioning by six potato cultivars grown on contaminated soil. International conference heavy metals in the environment CEP consultants, Edinburgh, pp 399–402

Harrison RM, Chirgawi MB (1989) The assessment of air and soil as contributors of some trace metals to vegetable plants. I. Use of a filtered air growth cabinet. Sci Total Environ 83:13–34

Haslett BS, Reid RJ, Rengel Z (2001) Zinc mobility in wheat: uptake and distribution of zinc applied to leaves or roots. Ann Bot 87:379–386

Hassler CS, Wilkinson KJ (2003) Failure of the biotic ligand and free-ion activity models to explain zinc bioaccumulation by *Chlorella kesslerii*. Environ Toxicol Chem 22:620–626

Hassler CS, Slaveykova VI, Wilkinson KJ (2004) Some fundamental (and often overlooked) considerations underlying the free ion activity and biotic ligand models. Environ Toxicol Chem 23:283–291

Hatch DJ, Jones LHP, Burau RG (1988) The effect of pH on the uptake of cadmium by four plant species grown in flowing solution culture. Plant Soil 105:121–126

Helal M, Baibagyshew E, Saber S (1998). Uptake of Cd and Ni by spinach, *Spinacea oleracea* (L.) from polluted soil under field conditions as affected by salt water irrigation. Agronomie 18:443–448

Hendrickson LL, Corey RB (1981) Effect of equilibrium metal concentrations on apparent selectivity coefficients of soil complexes. Soil Sci 131:163–171

Hopper JL, Parker DR (1999) Plant availability of selenite and selenate as influenced by the competing ions phosphate and sulfate. Plant Soil 210:199–207

Hough RL, Young SD, Crout NMJ (2003) Modelling of Cd, Cu, Ni, Pb and Zn uptake, by winter wheat and forage maize, from a sewage disposal farm. Soil Use Manage 19:19–27

Hough R, Breward N, Young SD, Crout NMJ, Tye AM, Moir AM, Thornton I, (2004) Assessing potential risk of heavy metal exposure from consumption of home-produced vegetables by urban populations. Environ Health Perspect 112 (2):215–221

Hough RL, Tye AM, Crout NMJ, McGrath SP, Zhang H, Young SD (2005) Evaluating a ‘Free Ion Activity Model’ applied to metal uptake by *Lolium perenne* L. grown in contaminated soils. Plant Soil 270:1–12

Hovmand MF, Tjell JC, Mosbaek H (1983) Plant uptake of airborne cadmium. Environ Pollut 30:27–38

Hudson RJM (2005) Trace metal uptake, natural organic matter, and the free-ion model. J Phycol 41:1–4

Ingwersen J, Streck JT (2006) Modeling the environmental fate of cadmium in a large wastewater irrigation area. J Environ Qual 35:1702–1714

Jansson G, Smolders E, Ruttens A, Römkens P, De Temmerman L, Bries J (2007) Teeltadvies voor de landbouw in kader van het Interreg project BeNeKempen. Final Report. OVAM, Mechelen, Belgium

Jones KC, Symon CJ, Johnston AE (1987) Long-term changes in soil and cereal grain cadmium: studies at rothamsted experimental station. Trace Subst Environ Health 21:450–461

Kalis EJJ, Temminghoff EJM, Weng L, van Riemsdijk WH (2006) Effects of humic acid and competing cations on metal uptake by *Lolium perenne*. Environ Toxicol Chem 25(3): 702–711

Khattak RA, Page AL, Parker DR, Bakhtan D (1991) Accumulation and interactions of arsenic, selenium, molybdenum and phosphorus in alfalfa. J Environ Qual 20:165–168

Krauss M, Wilcke W, Kobza J, Zech W (2002) Predicting heavy metal transfer from soil to plant: potential use of Freundlich-type functions. J Plant Nutr Soil Sci 165:3–8

Kuboi T, Noguchi A, Yazaki J (1986) Family-dependent cadmium accumulation characteristics in higher plants. Plant Soil 92:405–415

Lorenz SE, Hamon RE, McGrath SP, Holm PE, Christensen TH (1994) Applications of fertilizer cations affect cadmium and zinc concentrations in soil solutions and uptake by plants. Eur J Soil Sci 45:159–165

Luo CL, Shen ZG, Li XD (2005) Enhanced phytoextraction of Cu, Pb, Zn and Cd with EDTA and EDDS. Chemosphere 59:1–11

Maclean AJ (1974) Effects of soil properties and amendments on availability of zinc in soils. Can J Soil Sci 54:369–378

Maier NA, McLaughlin MJ, Heap M, Butt M, Smart MK (2002a) Effect of current-season application of calcitic lime and phosphorus fertilization on soil pH, potato growth, yield, dry matter content, and cadmium concentration. Commun Soil Sci Plant Anal 33:2145–2165

Maier NA, McLaughlin MJ, Heap M, Butt M, Smart MK (2002b) Effect of nitrogen source and calcitic lime on soil pH and potato yield, leaf chemical composition, and tuber cadmium concentrations. J Plant Nutr 25:523–544

Marin A, Masscheleyn P, Patrick W (1992) The influence of chemical form and concentration of arsenic on rice and tissue arsenic concentration. Plant Soil 139:175–183

Mattina MI, Isleyen M, Eitzer BD, Iannucci-Berger W, White JC (2006) Uptake by Cucurbitaceae of soil-borne contaminants depends upon plant genotype and pollutant properties. Environ Sci Technol 40:1814–1821

McBride MB (1989) Reactions controlling heavy metal solubility in soil. Adv Soil Sci 10:1–56

McGowen SL, Basta NT, Brown GW (2001) Use of diammonium phosphate to reduce heavy metal solubility and transport in smelter-contaminated soil. J Environ Qual 30:493–500

McGrath SP (1982) The uptake and translocation of tri- and hexa-valent chromium and effects on the growth of oat in flowing nutrient solution and in soil. New Phytol 92:381–390

McKenna IM, Chaney RL, Williams FM (1993) The effects of cadmium and zinc interactions on the accumulation and tissue distribution of zinc and cadmium in lettuce and spinach. Environ Pollut 79:113–120

McLaughlin MJ (2001a) Bioavailability of metals to terrestrial plants. In: Allen H (ed) Bioavailability of metals in terrestrial ecosystems: importance of partitioning for bioabailability to invertebrates, microbes and plants. Society of Environmental Toxicology and Chemistry (SETAC), USA, pp 39–68

McLaughlin MJ (2001b) Ageing of metals in soils changes bioavailability. Fact Sheet Environ Risk Assess 4:1–6. International council on metals and the environment, ON, Canada, September 2001

McLaughlin MJ (2002) Heavy metals. In: Lal R (ed) Encyclopedia of soil science. Marcel Dekker, New York

McLaughlin MJ, Maier NA, Rayment GE, Sparrow LABG, McKay A, Milham P, Merry RH, Smart MK (1997) Cadmium in Australian potato tubers and soils. J Environ Qual 26:1644–1649

McLaughlin MJ, Palmer LT, Tiller KG, Beech TA, Smart MK (1994a) Increased soil salinity causes elevated cadmium concentrations in field-grown potato tubers. J Environ Qual 23:1013–1018

McLaughlin MJ, Williams CMJ, McKay A, Gunton G, Jackson K, Dowling B, Kirkham R, Partington D, Smart MK, Tiller KG (1994b) Effect of cultivar on uptake of cadmium by potato tubers. Austr J Agric Res 45:1483–1495

McLaughlin MJ, Smolders E, Merckx R, Maes A (1997) Plant uptake of cadmium and zinc in chelator-buffered nutrient solution depends on ligand type. In: Ando T, Fujita K, Mae T, Matsumoto H, Mori S, Sekija J (eds) Plant nutrition for sustainable food production and environment. Kluwer Academic, Dordrecht, The Netherlands, pp. 113–118

McLaughlin MJ, Smolders E, Merckx R (1998) Soil-root interface: physicochemical processes. In: Huang PM et al. (eds) Soil chemistry and ecosystem health. SSSA Special Publication No. 52. Soil Science Society of America, Madison, Wisconsin, USA

Morel FMM (1983) Principles of aquatic chemistry. Wiley, New York, NY

Mosbaek H, Tjell JC, Hovmand MF (1989) Atmospheric lead input to agricultural crops in Denmark. Chemosphere 19:1787–1799

Nakamaru Y, Tagami K, Uchida S (2005) Distribution coefficient of selenium in Japanese agricultural soils. Chemosphere 58:1347–1354

Narwal RP, Singh BR (1998) Effect of organic material on partitioning, extractability and plant uptake of metals in an alum shale soil. Water Air Soil Pollut 103:405–421

Nolan AL, Lombi E, McLaughlin MJ (2003) Metal bioaccumulation and toxicity in soils – why bother with speciation? Austr J Chem 56:77–91

Nolan AL, Zhang H, McLaughlin MJ (2005) Prediction of zinc, cadmium, lead and copper bioavailability to wheat in contaminated soils using chemical speciation, diffusive gradients in thin films, extraction, and isotope dilution techniques. J Environ Qual 34:496–507

Page AL, Chang AC (1985) Fate of wastewater constituents in soil and groundwater: trace elements. In: Pettygrove GS, Asano T (eds) Lewis, USA, pp 13–1 to 13–16

Page V, Feller U (2005) Selective transport of zinc, manganese, nickel, cobalt and cadmium in the root system and transfer to the leaves in young wheat plants. Ann Bot 96:425–434

Plekhanova IO (2007) Transformation of Fe, Mn, Co, and Ni compounds in humic podzols at different moisture. Biol Bull 34:67–75

Parker DR, Chaney RL, Norvell WA (1995) Chemical equilibrium models: applications to plant nutrition research. In: Loeppert RH et al. (eds) Chemical equilibrium and reaction models. Soil science society of America special publication no. 42, SSSA, Madison, WI, USA, pp 163–200

Peijnenburg W, Baerselman R, de Groot A, Jager T, Leenders D, Posthuma L, Van Veen R (2000) Quantification of metal bioavailability for lettuce (*Lactuca sativa* L.) in field soils. Archiv Environ Contamin Toxicol 39:420–430

Plette ACC, Nederlof MM, Temminghoff EJM, van Riemsdijk WH (1999) Bioavailabilty of heavy metals in terrestrial and aquatic systems: a quantative approach. Environ Toxicol Chem 18:1882–1890

Popelka JC, Schubert S, Schulz R, Hansen AP (1996) Cadmium uptake and translocation during reproductive development of peanut. Angewandte Botanik 70:140–143

Preer JR, Abdi AN, Sekhon HS, Murchison GB (1995) Metals in urban gardens – effect of lime and sludge. J Environ Sci Health Part A 30:2041–2056

Preer JR, Sekhon HS, Stephens BR, Collins MS (1980) Factors affecting heavy-metal content of garden vegetables. Environ Pollut Series B-Chem Phys 1:95–104

Quaghebeur M, Rengel Z (2003) The distribution of arsenate and arsenite in shoots and roots of *Holcus lanatus* is influenced by arsenic tolerance and arsenate and phosphate supply. Plant Physiol 132:1600–1609

Reese RN, White CA, Winge DR (1992) Cadmium sulphide crystallites in Cd-(γ-EC)*n*G peptide complexes from tomato. Plant Physiol 98:225–229

Reid RJ, Dunbar KR, McLaughlin MJ (2003) Cadmium loading into potato tubers: the roles of the periderm, xylem and phloem. Plant Cell Environ 26:201–206

Rellan-Alvarez R, Ortega-Villasante C, Alvarez-Fernandez A, Del Campo FF, Hernandez LE (2006) Stress responses of Zea mays to cadmium and mercury. Plant Soil 279:41–50

Römkens PFAM, Guo HY, Chu CL, Liu TS, Chian CF, Koopmans GF (2009) Prediction of Cadmium uptake by brown rice and derivation of soil-plant transfer models to improve soil protection guidelines. Environ Poll 157:2435–2444

Römkens PFAM, Schuur GW, Lijzen JPA, Rietra RPJJ, Dirven-van Breemen EM (2005) Risico's van cadmium en lood in de Kempen. Alterra report nr 1129, Wageningen. www.alterra.wur.nl/UK/publications/Alterra+Reports/

Salt DE, Rauser WE (1995) MgATP-dependent transport of phytochelatins across the tonoplast of oat roots. Plant Physiol 107:1293–1301

Salt DE, Prince RC, Pickering IJ, Raskin I (1995) Mechanisms of cadmium mobility and accumulation in Indian mustard. Plant Physiol 109:1427–1433

Sauerbeck D, Styperek P (1985) Evaluation of chemical methods for assessing the Cd and Zn availability from different soils and sources. In: Leschber R et al. (eds) Chemical methods for assessing bioavailable metals in sludges and soils. ISBN 0853343594, Elsevier, London, pp 49–66

Sauvé S, Cook N, Hendershot WH, McBride MB (1996) Linking plant tissue concentrations and soil copper pools in urban contaminated soils. Environ Pollut 94(2):153–157

Sauvé S, Hendershot W, Allen HE (2000) Solid-solution partitioning of metals in contaminated soils: dependence on pH, total metal burden, and organic matter. Environ Sci Technnol 34:1125–1131

Shane BS, Littman CB, Essick LA, Gutenmann WH, Doss GJ, Lisk DJ (1988) Uptake of selenium and mutagens by vegetables grown in fly-ash containing greenhouse media. J Agric Food Chem 36:328–333

Sipter E, Rózsa E, Gruiz K, Tátrai E, Morvai V (2008) Site-specific risk assessment in contamined vegetable gardens. Chemosphere 71:1301–1307

Smolders E, McLaughlin MJ. 1996a. Effect of Cl and Cd uptake by Swiss chard in nutrient solution. Plant Soil 179:57–64

Smolders E, McLaughlin MJ (1996b) Chloride increases cadmium uptake in Swiss Chard in a resin-buffered nutrient solution. Soil Sci Soc Am J 60:1443–1447

Smolders E, Brans K, Foldi A, Merckx R (1999) Cadmium fixation in soils measured by isotopic dilution. Soil Sci Soc Am J 63:78–85

Sobotik M, Ivanov VB, Obroucheva NV, Seregin IV, Martin ML, Antipova OV, Bergmann H (1998) Barrier role of root system in lead-exposed plants. J Appl Bot 72:144–147

Sparrow LA, Salardini AA, Johnstone J (1994) Field studies of cadmium in potatoes (*Solanum tuberosum* L.). III. Response of cv. Russet Burbank to sources of banded potassium. Austr J Agric Res 45:243–249

Sposito E (1989) The chemistry of soils. Oxford University Press, New York, NY

Sposito G (1981) The operational definition of the zero point of charge in soils. Soil Sci Soc Am J 45:292–297

Sterckeman T, Perriguey J, Caël M, Schwartz C, Morel JL (2004) Applying a mechanistic model to cadmium uptake by *Zea mays* and *Thlaspi caerulescens*: consequences for the assessment of the soil quantity and capacity factors. Plant Soil 262(1–2):289–302

Tandy S, Schulin R, Nowack B (2006a) Uptake of metals during chelant-assisted phytoextraction with EDDS related to the solubilized metal concentration. Environ Sci Technol 40: 2753–2758

Tandy S, Schulin R, Nowack B (2006b) The influence of EDDS on the uptake of heavy metals in hydroponically grown sunflowers. Chemosphere 62:1454–1463

Thakali S, Allen H, Di Toro DM, Ponizovsky AA, Rooney CR, Zhao FJ, McGrath SP (2006) A terrestrial biotic ligand model. 1 Development and application to Cu and Ni toxicities to barley root elongation in soils. Environ Sci Technol 40:7085–7093

Thomas GM, Harrison HC (1989) Inheritance of cadmium concentration in lettuce. J Am Hortic Sci 114:121–125

Tiller KG, de Vries MPC, Spouncer LR, Smith L, Zarcinas B (1976) Environmental pollution of the Port Pirie region, 3. Metal contamination of home gardens in the city and their vegetable produce. 15. CSIRO Division of Soils Divisional Report, p 15

Tiller KG, Oliver DP, McLaughlin MJ, Merry RH, Naidu R (1997) Managing cadmium contamination of agricultural land. In: Iskandar IK, Adriano DC (eds) Remediation of soils contaminated with metals. Science Reviews, Northwood, pp 225–255

Tjell JC, Hovmand MF, Mosbaek H (1979) Atmospheric lead pollution of grass grown in a background area in Denmark. Nature 280:425–426

Traulsen BD, Schonhard G (1987) Amelioration of an old-copper-contaminated area model-experiment – area amelioration. Landwirtschaftliche Forschung 40:364–370

Tyler LD, McBride MB (1982) Influence of Ca, pH and humic acid on Cd uptake. Plant Soil 64:259–262

Vatamaniuk OK, Mari S, Lu YP, Rea PA (1999) AtPCS1, a pytochelatin synthase from Arabidopsis: isolation and in vitro reconstitution. Proc Natl Acad Sci USA 96:7110–7115

Verloo M, Willaert G (1988) De Accumulatie van cadmium door planten in relatie tot de fysicochemische eigenschappen van de grond. Cadmium, Voorkimen, Impaket en Sanering, Symposiu, Genk

Vogeli-Lange R, Wagner GJ (1990) Subcellular Localization of cadmium and cadmium-binding peptides in tobacco leaves. Plant Physiol 92:1086–1093

Voutsa D, Grimanis A, Samara C (1996) Trace elements in vegetables grown in an industrial area in relation to soil and air particulate matter. Environ Pollut 94:325–335

de Vries W, Romkens PFAM, Schutze G (2007) Critical soil concentrations of cadmium, lead, and mercury in view of health effects on humans and animals. Rev Environ Contamin Toxicol 191:91–130

de Vries W, Romkens PFAM, Bonten LTC (2008) Spatially explicit integrated risk assessment of present soil concentrations of cadmium, lead, copper and zinc in the Netherlands. Water Air Soil Pollut 191(1–4):199–215

Wang JR, Zhao FJ, Meharg AA, Raab A, Feldmann J, McGrath SP (2002) Mechanisms of arsenic hyperaccumulation in Pteris vittata. Uptake kinetics, interactions with phosphate, and arsenic speciation. Plant Physiol 130:1552–1561

Wang JL, Fang W, Yang ZY, Yuan JG, Zhu Y, Yu H (2007) Inter- and intraspecific variations of cadmium accumulation of 13 leafy vegetable species in a greenhouse experiment. J Agric Food Chem 55:9118–9123

Wang XP, Shan XQ, Zhang SZ, Wen B (2004) A model for evaluation of the phytoavailability of trace elements to vegetables under the field conditions. Chemosphere 55:811–822

Warren GP, Alloway BJ, Lepp NW, Singh B, Bochereau FJM, Penny C (2003) Field trials to assess the uptake of arsenic by vegetables from contaminated soils and soil remediation with iron oxides. Sci Total Environ 311:19–33

Weng LP, Lexmond TM, Wolthoorn A, Temminghoff EJM, van Riemsdijk WH (2003) Phytotoxicity and bioavailability of nickel: chemical speciation and bioaccumulation. Environ Toxicol Chem 22(9): 2180–2187

Weng LP, Wolthoorn A, Lexmond TM, Temminghoff EJM, van Riemsdijk WH (2004) Understanding the effects of soil characteristics on phytotoxicity and bioavailability of nickel using speciation models. Environ Sci Technol 38:156–162

Whiting SN, Broadley MR, White PJ (2003) Applying a solute transfer model to phytoextraction: Zinc acquisition by *Thlaspi caerulescens*. Plant Soil 249:45–56

Williams CH, David DJ (1973) The effect of superphosphate on the cadmium content of soils and plants. Austr J Soil Res 11:43–56

Wong MTF, Wittwer K (2009) Positive charge discovered across Western Australian wheatbelt soils challenges key soil and nitrogen management assumptions. Austr J Soil Res 47:127–135

Woolson EA (1973) Arsenic phytotoxicity and uptake in 6 vegetable crops. Weed Sci 21:524–527

Xu JL, Thornton I (1985) Arsenic in garden soils and vegetable crops in Cornwall, England – implications for human health. Environ Geochem Health 7:131–133

Xue Q, Harrison HC (1991) Effects of soil zinc, pH, and cultivar on cadmium uptake in leaf lettuce (*Lactuca sativa* L. var. crispa). Commun Soil Sci Plant Anal 22:975–991

Yuran GT, Harrison HC (1986) Effects of genotype and sewage sludge on cadmium concentration in lettuce leaf tissue. J Am Hortic Sci 111:491–494

Zhao FJ, Adams ML, Dumont C, McGrath SP, Chaudri AM, Nicholson FA, Chambers BJ, Sinclair AH (2004) Factors affecting the concentrations of lead in British wheat and barley grain. Environ Pollut 131:461–468

Chapter 9
Uptake of Organic Contaminants from Soil into Vegetables and Fruits

Stefan Trapp and Charlotte N. Legind

Abstract Contaminants may enter vegetables and fruits by several pathways: by uptake with soil pore water, by diffusion from soil or air, by deposition of soil or airborne particles, or by direct application. The contaminant-specific and plant-specific properties that determine the importance of these pathways are described in this chapter. A variety of models have been developed, specific for crop types and with steady-state or dynamic solutions. Model simulations can identify sensitive properties and relevant processes. Persistent, polar (log $K_{OW} < 3$) and non-volatile ($K_{AW} < 10^{-6}$) contaminants have the highest potential for accumulation from soil, and concentrations in leaves may be several hundred times higher than in soil. However, for most contaminants the accumulation in vegetables or fruits is much lower. Lipophilic (log $K_{OW} > 3$) contaminants are mainly transported to leaves by attached soil particles, or from air. Volatile contaminants have a low potential for accumulation because they quickly escape to air. Experimental data are listed that support these model predictions, but underline also the high variability of accumulation under field conditions. Plant uptake predictions are uncertain, due to the immense variation in environmental and plant physiological conditions. Uptake of organic contaminants into vegetables and fruits may lead to human health risks, but it may also be used to delineate subsurface plumes and monitor Natural Attenuation. Most models mentioned in this chapter are freely available from the authors.

Contents

S. Trapp (✉)
Department of Environmental Engineering, Technical University
of Denmark, Lyngby, Copenhagen, Denmark
e-mail: sttr@env.dtu.dk

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_9, © Springer Science+Business Media B.V. 2011

9.1 Introduction

A major concern of European citizens is residues of pesticides in fruits and vegetables (EFSA 2006). But not just pesticides pose a risk. In fact, for a whole range of organic contaminants, uptake via diet is often the primary contribution to human exposure (SCF 2002; Travis and Hattemer-Frey 1991). Contaminants may be taken up into plants and subsequently accumulate in the human food chain (Czub and McLachlan 2004) and affect the health of humans. Consequently, uptake of contaminants into plants is an essential part of most exposure models, for example CSOIL (Brand et al. 2007), CLEA (DEFRA 2002) and EUSES (EC 2003).

About 20,000 plant species are used by the human race, and about 600 species are cultivated (Franke 1987). Additionally, about 250,000 wild plants grow on earth (Sitte et al. 1991). Thus, there is a large variability in plant properties. Also the growth conditions vary, depending on soil type, soil properties, climatic conditions and agricultural practice. Similarly, the number of organic contaminants is very high. More than 5 million compounds have been synthesized. Around 30,000 compounds are marketed in Europe, and contaminants released to the environment may also be metabolized. Thus, from a researcher's point of view, the number of possible combinations of plant species, contaminants and environmental conditions is close to infinite. Nevertheless, general patterns are known and process-oriented models have been established.

This chapter will give an overview of uptake processes of organic contaminants from soil into plants, on prediction methods and on experimental results. Model simulations will be carried out to identify the chemical properties that control the accumulation in food crops. These predictions will be compared with experimental results, in order to determine the potential of soil contaminants for accumulation in food crops. For the calculation of exposure through vegetable consumption, see Chapter 11 by Elert et al., of this book.

9.2 Uptake and Transport Processes

Contaminants in the environment can enter plants by various ways (Fig. 9.1). The main passive transport and uptake processes from soil are:

- uptake with transpiration water;
- diffusion from soil into roots;
- attachment of soil particles, eventually followed by diffusion into plant tissue.

However, contaminants can also be present in air. The main uptake processes from air are:

- diffusive (gaseous) exchange with air;
- wet and dry particle deposition from air on plant surfaces followed by diffusion into plant tissue.

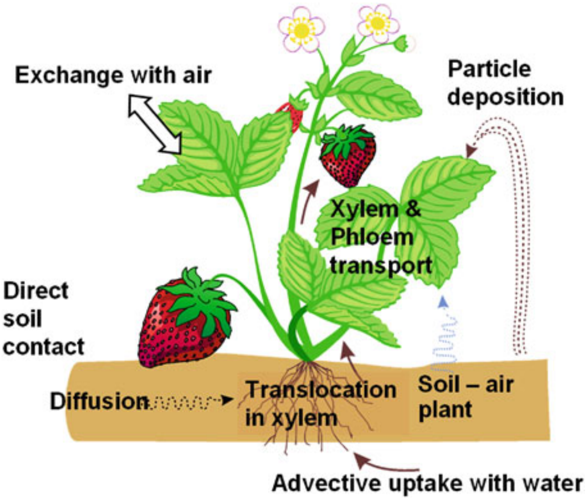


Fig. 9.1 Transport and uptake processes in the soil-air-plant system

Inside the plant, the phloem and xylem flux may distribute the contaminants. The xylem sap flows from the roots into the stem to the leaves and, to some extent, to fruits. The phloem sap flows from the leaves to all growing parts of the plant and to fruits and storage organs, such as tubers. The relative importance of these processes varies with plant type, environmental conditions and properties of contaminants. Active uptake processes, which involve energy or enzymes of the plant, may also play a role, but have not yet been shown to be of large relevance for environmental organic contaminants.

9.3 Empirical Methods for Estimating Uptake of Contaminants into Plants

Laboratory and field experiments have been conducted to determine the uptake of contaminants from soil into plants. Single bioconcentration factors for contaminants as well as regressions were established for predicting the uptake of organic contaminants from soil into plants.

9.3.1 Bioconcentration Factors

The ratio of contaminant concentration in an organism to contaminant concentration in the surrounding medium is called the bioconcentration factor (*BCF*). Measurements of concentrations in plant tissues and concentrations in soil will yield a *BCF* plant to soil, which is defined as

$$BCF = \frac{C_{\text{plant}}}{C_{\text{soil}}} \tag{9.1}$$

where C_{Plant} is the concentration in plant tissues and C_{Soil} is the concentration in soil (ideally at steady state, but practically at harvest). This BCF will only apply to the specific contaminant and soil type used for the determination.

Care must be taken in cases where a measurable background concentration in plants is present. Because then, for low soil concentration ($C_{\text{Soil}} \rightarrow 0$), the concentration ratio *BCF* can be very high ($C_{\text{Plant}}/C_{\text{Soil}} \rightarrow \infty$). For higher soil concentrations, however, the *BCF* decreases and approaches a constant value. This pattern was occasionally interpreted as a variable *BCF* with soil concentration, i.e. a decreasing *BCF* with increasing soil concentration. A real-world example is the ratio between the measured concentration of p,p′-DDT in radishes and in soil. The concentration ratio is high at low soil concentrations, and decreases for higher soil concentrations. A plausible explanation for this pattern is that plants have a limited sorption capacity for organic contaminants, which becomes saturated at higher soil concentrations. However, a more likely interpretation is that the uptake into plants is from two different and independent sources, namely from soil and from air. When soil concentrations are very low there still is a background contamination of the plant tissue originating from air (Mikes et al. 2009).

Instead of simply calculating the concentration ratio of plant to soil, the relationship between concentrations in plant and soil can be quantified by a linear regression between both if measurements at different concentration levels are available. The slope of the regression between soil concentration as predictor variable and plant concentration as estimated variable can be interpreted as the *BCF* plant to soil, while the *y*-axis-intercept can be interpreted as the constant background concentration due to uptake from air.

$$C_{\text{Plant}} = BCF \times C_{\text{Soil}} + C_{\text{Background}} \tag{9.2}$$

where $C_{\text{Background}}$ is the constant concentration due to uptake from air.

This method has several advantages:

- all measured values contribute to the calculated *BCF*;
- variations in the measured concentrations are adequately considered;
- the y-axis gives the concentration in plants due to the (constant) concentration in air;
- the square of the correlation coefficient (r^2) describes how much of the variance in the measured concentration in plants is explained by the variance of the concentrations in soil.

An example is shown in Fig. 9.2. It shows the *BCF*s for p,p′-DDT in radishes which decrease with increasing soil concentration (Mikes et al. 2009), but with C_{plant} plotted as a function of C_{Soil}. The slope of the regression curve, i.e. the value 0.17, is the *BCF* derived from all measured values minus the background concentration in air. The *BCF* is statistically highly significant. The explained variance r^2 is 0.98, which means that the increase of concentrations in plants can be explained almost completely by the increase of concentration in soil.

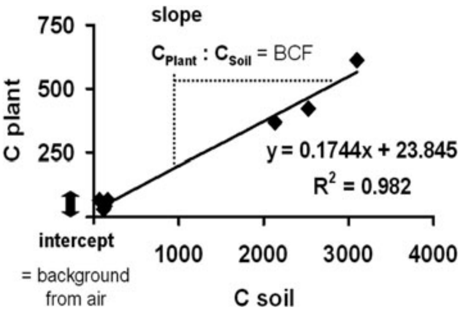


Fig. 9.2 Calculation of *BCF* from the slope of the regression between concentrations in soil and plant root; calculation of background concentration due to uptake from air from the y-axis intercept of the regression. Example from (Mikes et al. 2009), p,p′-DDT in roots and soil

9.3.2 Regression Equations

*BCF*s for neutral organic contaminants are related to contaminant properties. This is applied in regression equations where physico-chemical properties like lipophilicity of the contaminants are correlated to the *BCF*s. These regressions based on measured data allow us to estimate concentrations of contaminants in plant tissue from the octanol-water partitioning coefficient, K_{OW}, in case of uptake from soil (e.g. (Travis and Arms 1988)) or from the octanol-air partitioning coefficient, K_{OA}, in case of uptake from air (Kömp and McLachlan 1997).

A frequently applied regression for the uptake of organic contaminants from soil into above-ground plants stems from Travis and Arms (1988):

$$\log BCF_V = 1.588 - 0.578 \times \log K_{OW}\ (n = 29, r^2 = 0.53) \tag{9.3}$$

where BCF_v is the bioconcentration factor of vegetation to soil (kg $\mathrm{kg_{dw}}^{-1}$) and K_{OW} (L L^{-1}) expresses the lipophilicity of the contaminant. The curve was fitted with 29 data points in the range $1.15 \leq \log K_{OW} \leq 9.35$. A limitation of the regression is that the uptake from air remains unknown. This may lead to false predictions.

9.3.3 Root Concentration Factor

Phase equilibrium is the endpoint of diffusion and is achieved when the activity of the contaminant in the root tissue is equal to the activity of the contaminant in the external solution (Lewis 1907). The concentration ratio between root and aqueous solution in phase equilibrium is called the *root concentration factor RCF* (L kg^{-1}) (Shone and Wood 1974).

$$RCF = \frac{\text{Concentration in root (mg g}^{-1}\text{)}}{\text{Concentration in solution (mg mL}^{-1}\text{)}} \tag{9.4}$$

The most widely applied regression for the *RCF* is from Briggs et al. (1982). It has been fitted to data derived from experiments with intact and macerated barley roots. For neutral organic chemicals (*phenylureas* and *o-methylcarbamoyloximes*), the relation between log K_{OW} and sorption to roots was:

$$\log(RCF - 0.82) = 0.77 \log K_{ow} - 1.52 \ (n = 7, r^2 = 0.96). \tag{9.5}$$

The dependency on the log K_{OW} was explained by lipophilic sorption of the contaminants to plant lipids. The value of 0.82 was interpreted as water content of the roots. A similar result was obtained for cut pieces of bean roots and stems for *N-methyl-arylcarbamates* (Trapp and Pussemier 1991):

$$\log(RCF - 0.85) = 0.557 \log K_{ow} - 1.34 \ (n = 12, r^2 = 0.92). \tag{9.6}$$

This equation gives lower root concentration factors for lipophilic contaminants (Fig. 9.3). Both *RCF*-regressions describe partitioning to water. But roots typically grow in soil. The sorption of organic contaminants to soil is related to the organic carbon content, *OC*. Many regressions for the organic carbon normalised sorption to soil, K_{OC} (L $\text{kg}_{\text{dw}}{}^{-1}$), were established, for example for neutral organic contaminants (EC 2003)

$$\log K_{OC} = 0.81 \log K_{OW} + 0.1 \ (n = 81, r^2 = 0.89) \tag{9.7}$$

Bulk soil consists of solids, water and air. For the concentration ratio between bulk soil and soil pore water, K_{SW} (L kg^{-1}) follows

$$K_{SW} = \frac{C_{\text{Soil}}}{C_W} = \frac{OC \times K_{OC} \times \rho_{S,dry} + \theta \times \rho_W}{\rho_{S,wet}} = \frac{1}{K_{SW}} \tag{9.8}$$

where *OC* is the fraction of organic carbon in soil (kg kg^{-1}), θ is the water content of soil (L kg^{-1}), ρ_W is the density of water (1 kg L^{-1}), $\rho_{S,dry}$ is the soil dry density (kg_{dw} L^{-1}) and $\rho_{S,wet}$ is the soil wet density (kg_{ww} L^{-1}). Division of *RCF* with K_{SW} gives the equilibrium concentration of the ratio of roots to soil.

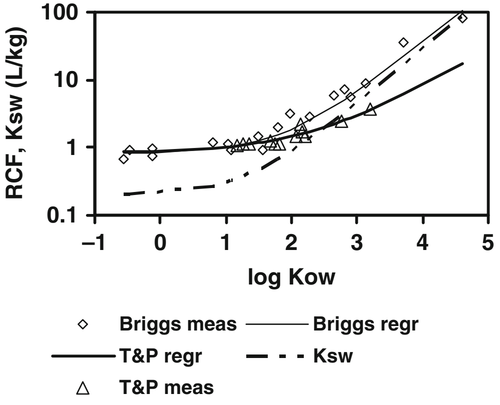


Fig. 9.3 Measured data and root concentration factor (*RCF*)-regression lines of Briggs et al. (1982) and Trapp and Pussemier (1991), compared to the concentration ratio bulk soil to water (K_{SW})

Both *RCF*-regressions (Eqs. 9.5 and 9.6), the experimental data of Briggs et al. (1982) and the concentration ratio between bulk soil and soil pore water, K_{SW}, for a typical soil ($OC = 0.025$ kg kg^{-1} and $W = 0.2$ kg kg^{-1}) are plotted in Fig. 9.3. For low K_{OW} values *RCF* is higher than K_{SW}, due to the higher water content of roots. For higher K_{OW} values Briggs' *RCF* regression and the K_{SW} equation (Eq. 9.8) yield similar results. This suggests that the sorption capacity of roots equals that of soil, because the content of organic carbon in soil (in this case 2.5%) is similar to the lipid content of roots (about 2–3%, including waxes and lignin), and the slope of the log K_{OW} in the regressions is similar (0.81 for K_{OC} in Eq. 9.7 and 0.77 for *RCF* in Eq. 9.5).

9.3.4 Partition Coefficients for Stem and Leaves

Briggs et al. (1983) measured the sorption to macerated barley stems and predicted $K_{\text{stem/xylem sap}}$ (L kg^{-1}), which is the concentration of contaminants in stem tissue divided by the concentration in xylem sap, related to the log K_{OW} of the contaminants:

$$\log(K_{\text{stem/xylemsap}} - 0.82) = 0.95 \log K_{OW} - 2.05 \ (n = 8, r^2 = 0.96). \tag{9.9}$$

Trapp et al. (1994) interpreted the regressions derived for sorption to roots and stems as equilibrium partition coefficients between plant tissue and water, K_{PW} (L kg^{-1}), and introduced the general equation:

$$K_{PW} = W + LaK_{OW}^{b} \tag{9.10}$$

where W (L kg^{-1}) and L (kg kg^{-1}) are water and lipid content of the plant, b is a correction factor for differences between solubility in octanol and sorption to plant lipids (in the regressions of Briggs et al. (1982, 1983) b was 0.77 for roots and 0.95 for leaves), and a is a factor correcting density differences between water and n-octanol ($1/\rho_{\text{Octanol}} = 1.22$ L kg^{-1}, where ρ_{Octanol} is the density of octanol). When parameterized accordingly, this equilibrium approach gives the same results for roots as the Briggs *RCF*-regression.

Stems and leaves are in contact with air. The sorption equilibrium of contaminants between leaves and air can be described as follows:

$$K_{LA} = C_L/C_A = K_{LW}/K_{AW} \tag{9.11}$$

where K_{LA} is the partition coefficient between leaves and air (L kg^{-1}), K_{LW} (L kg^{-1}) is the partition coefficient between leaves and water (Eq. 9.10) and K_{AW} (L L^{-1}) is the partition coefficient between air and water (also known as the dimensionless Henry's Law constant). Instead of estimating K_{LA} from K_{OW} and K_{AW}, K_{LA} was often directly fitted to K_{OA}, i.e. the partition coefficient between octanol and air (e.g., Kömp and McLachlan 1997).

9.3.5 Translocation from Roots into Stem and Leaves

Translocation of contaminants from roots into stems in the xylem is often described by the transpiration stream concentration factor *TSCF* (Russell and Shorrocks 1959):

$$TSCF = \frac{\text{Concentration in xylem sap (mg mL}^{-1}\text{)}}{\text{Concentration in solution (mg mL}^{-1}\text{)}} \tag{9.12}$$

From experimental data, the concentration in xylem sap is found from the mass of contaminant in shoots divided by the volume of transpired water. Briggs et al. (1982) fitted a Gaussian optimum curve to their data:

$$TSCF = 0.784 \times e^{-(\log K_{OW} - 1.78)^2/2.44} \; (n = 17, r^2 = 0.73) \tag{9.13}$$

Later experiments by other research groups yielded similar results (Burken and Schnoor 1998; Hsu et al. 1990; Sicbaldi et al. 1997). However, recent research has shown that these regressions might not be valid for the polar contaminants ($\log K_{OW} < 1$) and a sigmoidal relationship between *TSCF* and $\log K_{OW}$ has been established (Dettenmaier et al. 2009):

$$TSCF = \frac{11}{11 - 2.6^{\log K_{OW}}} \; (n = 118, \; r^2 = 0.68) \tag{9.14}$$

9.4 Mechanistic Models for Estimating Uptake of Contaminants into Plants

The first author met D. Mackay at his talk at the University of Bayreuth in 1986, where he encouraged his audience to develop mechanistic models for the plant uptake of organic contaminants. So we did (Trapp et al. 1990). D. Mackay and his team also developed some of the early models for this purpose, formulated with the fugacity approach (Paterson et al. 1994). These models were later simplified by Hung and Mackay (1997) to three mass balance equations and solved numerically. At about the same time, the numerical model PlantX was developed (Trapp et al. 1994). Later, crop-specific models were derived, i.e. specific models for roots (Trapp 2002), potatoes (Trapp et al. 2007a), leaves (Trapp and Matthies 1995) and fruits (Trapp 2007). These models are all based on the same physico-chemical principles and describe the same basic processes, such as advective uptake into plants, diffusive uptake, chemical equilibrium, transport in xylem and phloem, dilution by growth, and particle deposition from soil and air. The actually occurring processes and their parameterization depend on the type of crop (Fig. 9.4). Plant models were also developed by other groups, for example a partition-limited model by Chiou et al. (2001) and the models for pesticide uptake by Fujisawa et al. (2002a, b). For

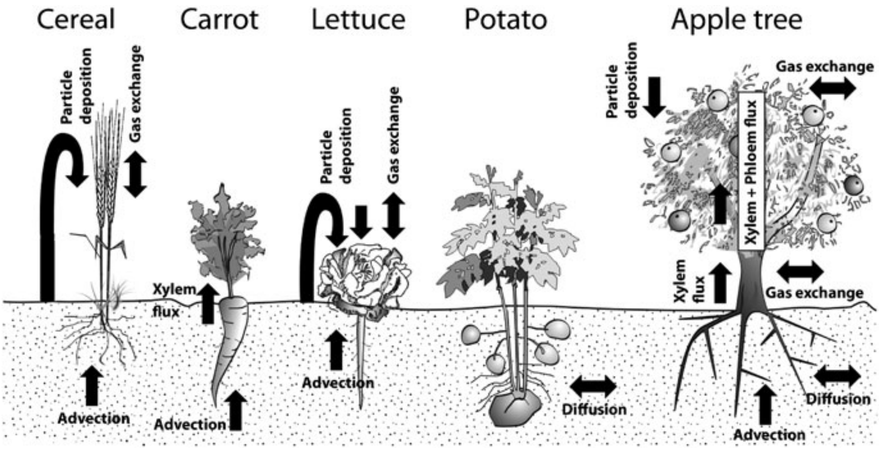


Fig. 9.4 Overview of crop-specific plant uptake processes (Legind and Trapp 2009)

electrolytes, such as acids and bases, very different approaches were developed. Examples of such models are the phloem transport model by Kleier (1988), the model relating to pesticide spray application by Satchivi (2000a, b) and the cell model (Trapp 2000, 2004).

9.4.1 Processes to Include in a Plant Uptake Model

A plant uptake model should consider both uptake from soil and air into plants. A relatively simple model includes the compartments soil, roots and leaves (or fruits or grains) and the processes:

- continuous and pulse input to all compartments (soil, roots and leaves);
- degradation, leaching, run-off and plant uptake, resulting in loss from soil;
- uptake into roots with the transpiration water;
- growth dilution, degradation and metabolism in roots;
- translocation from roots to leaves (or fruits) with the transpiration stream;
- loss from leaves to air;
- deposition from air to leaves;
- transport to leaves with attached soil;
- growth dilution, degradation and metabolism in leaves.

9.4.2 Mass Balance for a Dynamic Plant Uptake Model

Though not in the scope of this chapter, a mass balance for soil can be described as follows.

Change of contaminant mass in soil is deposition from air minus leaching, run-off, volatilization, degradation and uptake into roots. Division by soil mass, M_S, results in the concentration in soil:

$$\frac{dC_s}{dt} = \frac{I_s}{M_s} - k_s \times C_s \tag{9.15}$$

where I_S (mg d^{-1}) is input to soil (including deposition from air), and k_S (d^{-1}) is the sum of all first-order loss rates from soil. For parameters for the equation, see other sections of this book.

The mass balance for thick roots, such as carrots, can be described as follows.

Change of contaminant mass in roots is influx with water minus outflux with xylem sap. Diffusive uptake is not considered, since it only makes a small change in the concentration in roots. The root is described with the following equation:

$$\frac{dm_R}{dt} = Q \times C_W - Q \times C_{Xy} \tag{9.16}$$

where m_R is the mass of contaminant in roots (mg), Q is the transpiration stream (L d^{-1}), C_W is the concentration in soil pore water (mg L^{-1}) and C_{Xy} is the concentration in the xylem at the outflow of the root (mg L^{-1}). If the xylem sap is in equilibrium with the root, the concentration is $C_{Xy} = C_R/K_{RW}$. K_{RW} (L kg^{-1}) is the partition coefficient between root and water (Eq. 9.10). The concentration in soil pore water, C_W, is $C_S \times K_{WS}$. Substituting these expressions in Eq. 9.16 gives the following equation:

$$\frac{d(C_R \times M_R)}{dt} = \frac{dm_R}{dt} = Q \times K_{WS} \times C_S - \frac{Q}{K_{RW}} \times C_R \tag{9.17}$$

If plant growth is exponential, and the ratio Q/M_R (transpiration stream Q to root mass M_R (kg)) is constant, the growth by exponential dilution can be considered by a first-order growth rate k_R (d^{-1}). If first order degradation or metabolism occurs, the rate k_R is the sum of the loss processes and the growth dilution. Division by mass of the root results in the concentration in roots:

$$\frac{dC_R}{dt} = \frac{Q}{M_R} \times K_{WS} \times C_S - \frac{Q}{M_R \times K_{RW}} \times C_R - K_R \times C_R \tag{9.18}$$

The mass balance for leaves can be described as follows.

Change of contaminant mass in leaves is influx with transpiration water plus gaseous and particulate deposition from air minus diffusion to air. This results in the following equation:

$$\begin{aligned}\frac{dm_L}{dt} =& \frac{Q}{K_{RW}} C_R + A_L \times g_L \times (1 - f_P) \times C_A + \frac{A_L \times v_{dep}}{2} \times f_P \times C_A \\ &- \frac{A_L \times g_L \times 1000\, Lm^{-3}}{K_{LA}} \times C_L\end{aligned} \tag{9.19}$$

where A_L is leaf area (m^2), K_{LA} is the partition coefficient between leaves and air (L kg^{-1}), C_A is the total concentration in air (mg m^{-3}) and f_P (−) is the fraction of the total concentration in air that is adsorbed on particles. Uptake from air can either be by diffusive exchange in the gas phase with conductance g_L (m d^{-1}), or by deposition of particles on the surface of the leaves ($A_L/2$) with velocity v_{dep} (m d^{-1}). The concentration in leaves is as follows:

$$\frac{dC_L}{dt} = \frac{Q}{M_L \times K_{RW}} \times C_R + \frac{A_L \times g}{M_L} \times (1 - f_P) \times C_A + \frac{A_L \times v_{dep}}{2 \times M_L} \times f_P \times C_A - \frac{A_L \times g_L \times 1000\,Lm^{-3}}{K_{LA} \times M_L} \times C_L - k_L \times C_L \tag{9.20}$$

where k_L (d^{-1}) again is the first-order rate that includes growth dilution and biotic and abiotic (photolysis) degradation processes. The first term of the equation quantifies translocation from roots to leaves and replaces the *TSCF* in earlier model versions (Trapp and Matthies 1995). The advantages of this new formulation are as follows:

- There is a relation between concentrations in roots and in leaves. This allows, for example, calculation of the fate of metabolites formed in roots.
- The *TSCF* is related to plant physiological parameters, such as transpiration Q, growth rate k and partitioning between root tissue and xylem, K_{RW} (Eq. 9.10).

The calculated concentration ratio between the xylem and the external solution is close to the calculated concentration resulting from the empirical *TSCF*-regression by Dettenmaier et al. (2009) for all contaminants, and to the *TSCF*-regressions following a Gaussian curve for contaminants with log $K_{OW} > 2$ (Trapp 2007). Trapp (2007) speculates, based on this equation, that plants growing in soil outdoors would have a different *TSCF*-curve than plants grown in hydroponic solutions. This is due to the formation of root hairs in soil, which leads to better diffusive uptake of polar contaminants and subsequently higher *TSCF*-values (the *TSCF* remains high (i.e. near 1) for contaminants with a log $K_{OW} < 1$). Dettenmaier et al. (2009) suggests that differences in experimental methods and plant growth conditions cause the disparity.

An additional process not considered in Eq. 9.20 is the contamination of leaves with attached soil, R (kg soil kg plant^{-1} (wet weight)). A convenient way of calculation is to add the concentration due to attachment of particles from soil with subsequent deposition on leaves to the calculated C_L, as follows:

$$C_{L,\text{Final}} = C_{L,\text{Calc}} + R \times C_{\text{Soil}} \tag{9.21}$$

Default values for R range from 0.001 kg kg^{-1} to 0.01 kg kg^{-1} (see Section 9.6.3).

Mass balances for fruit and grain can be set up analogous to Eq. (9.20). However, the processes and parameters differ. For grain, we assume no particle deposition and a lower soil attachment value.

9.4.3 Steady-State Solution for the Root and Leaf Model

The steady-state ($t \to \infty$) concentration in roots, C_R, with constant concentration in soil, C_{Soil}, is as follows:

$$C_R = \frac{Q}{\frac{Q}{K_{RW}} + K_R \times M_R} \times K_{WS} \times C_{Soil} \tag{9.22}$$

For leaves, the steady state concentration is:

$$C_L = \frac{I}{a} \tag{9.23}$$

where I is the sum of all input terms (mg kg^{-1} d^{-1}):

$$I = \frac{Q}{M_L \times K_{RW}} \times C_R + \frac{A_L \times g_L}{M_L} \times (1 - f_P) \times C_A + \frac{A_L \times v_{dep}}{2M_L} \times f_p \times C_A \tag{9.24}$$

and a is the sum of all loss processes (d^{-1}):

$$a = \frac{A_L \times g_L \times 1000\ Lm^{-3}}{K_{LA} \times M_L} + k_L \tag{9.25}$$

9.4.4 General Solutions for a Cascade Model

The system of three linear differential equations (Eqs. 9.15, 9.18 and 9.20) can be solved analytically or numerically for continuous or pulse input. Continuous input occurs from atmospheric deposition to soil and leaves, whereas pulse inputs vary, e.g. inputs from accidents, pesticide spray application and application of manure or compost.

The differential equations for the contaminant concentration in soil, root and leaves can be treated as a diagonal matrix, so that:

$$\frac{dC_1}{dt} = -k_1 C_1 + I_1/M_1 \tag{9.26}$$

$$\frac{dC_2}{dt} = +k_{12} C_1 - k_2 C_2 + I_2/M_2 \tag{9.27}$$

$$\frac{dC_3}{dt} = +k_{23} C_2 - k_3 C_3 + I_3/M_3 \tag{9.28}$$

where indices 1–3 refer to soil, roots and leaves, respectively; C (mg kg^{-1}) is concentration; k_1, k_2 and k_3 are the sum of all first-order loss processes in compartment 1, 2 and 3, respectively, and k_{12} and k_{23} are the transfer rates from compartment 1 to 2 and 2 to 3, respectively. I_i (mg d^{-1}) describes the constant input to the compartments, e.g., from air, and M_i (kg) is the mass of compartment i, $i = 1, 2, 3$. The matrix elements k and I can be derived from the differential equations above (Eqs. 9.15, 9.18 and 9.20).

Linear differential equations approach steady state for $t \to \infty$, i.e. the change of concentration with time is zero, $dC/dt = 0$. The steady-state solutions for matrix equations 1 (soil), 2 (roots) and 3 (leaves) with continuous input are as follows:

$$C_1(t \to \infty) = \frac{I_1}{k_1 M_1} \tag{9.29}$$

$$C_2(t \to \infty) = \frac{I_2}{k_2 M_2} + \frac{k_{12}}{k_2}\, C_1\,(t \to \infty) \tag{9.30}$$

$$C_3(t \to \infty) = \frac{I_3}{k_3 M_3} + \frac{k_{23}}{k_3}\, C_2\,(t \to \infty) \tag{9.31}$$

The steady-state solution follows the general scheme:

$$C_n(t \to \infty) = \frac{I_n}{k_n M_n} + \frac{k_{n-1,n}}{k_n} \times C_{n-1}(t \to \infty) \tag{9.32}$$

where n is the compartment number.

The analytical solutions for the differential equations 1 (soil), 2 (roots) and 3 (leaves) for a pulse input is the same as for initial concentrations $C(0) \neq 0$:

$$C_1(t) = C_1(0) \times e^{-k1t} \tag{9.33}$$

$$C_2(t) = k_{12} C_1(0) \times \left(\frac{e^{-k_1 t}}{(k_2 - k_1)} + \frac{e^{-k_2 t}}{(k_1 - k_2)} \right) + C_2(0) \times e^{-k_2 t} \tag{9.34}$$

$$\begin{aligned} C_3(t) = {} & k_{12} k_{23} C_1(0) \times \left\{ \frac{e^{-k_1 t}}{(k_1 - k_2)(k_1 - k_3)} + \frac{e^{-k_2 t}}{(k_2 - k_1)(k_2 - k_3)} + \frac{e^{-k_3 t}}{(k_3 - k_1)(k_3 - k_2)} \right\} \\ & + k_{23} C_2(0) \times \left(\frac{e^{-k_2 t}}{(k_3 - k_2)} + \frac{e^{-k_3 t}}{(k_2 - k_3)} \right) + C_3(0) \times e^{-k_3 t} \end{aligned} \tag{9.35}$$

The general solution scheme for pulse input to soil only, i.e. $C_1(0) \neq 0$ and $C_n(0) = 0$ with $n \geq 2$ is as follows:

$$C_n(t) = \prod_{i-1}^{n-1} k_{i,i+1} C_1(0) \times \left\{ \sum_{j=1}^{n} \frac{e^{-k_j t}}{\prod\limits_{k=1,k \neq j}^{n} (k_k - k_j)} \right\} \tag{9.36}$$

and for pulse input into all compartments, i.e. $C_n(0) \neq 0$ with $n \geq 1$:

$$C_n(t) = \sum_{a=1}^{n-1} \left(C_a(0) \prod_{i=1}^{n-1} k_{i,i+1} \times \sum_{j=a}^{n} \frac{e^{-k_j t}}{\prod\limits_{k=a,k\neq j}^{n} (k_k - k_j)} \right) + C_n(0) \times e^{-k_n t} \quad (9.37)$$

This solution scheme can offer solutions for several soil-plant-air cascade system variations. For example, in analogy with this example, compartment 1 could be soil, compartment 2 roots, and compartment 3 leaves. Or, alternatively, compartment 1 could be parent contaminant in soil, compartment 2 metabolite in soil, compartment 3 metabolite in root, compartment 4 metabolite in leaves, and compartment 5 metabolite of metabolite in leaves.

Also for repeated applications there is a solution. This is a situation in which one, two or more subsequent pulse inputs occur. In that case, the resulting concentration can be calculated by adding the concentrations resulting from steady state and one, two or more pulse inputs. To this purpose, the simulation is split up into several periods. The concentration vector ***C(t)*** at the end of a specific period serves as initial

Table 9.1 Input data set for the standard model for the calculation of plant uptake (normalised to 1 m^2 of soil)

Parameter	Symbol	Value	Unit
Soil			
Soil wet density	ρ_{wet}	1.95	$\text{kg}_{\text{ww}}\ \text{L}^{-1}$
Organic carbon content	OC	0.02	$\text{kg}\ \text{kg}_{\text{ww}}{}^{-1}$
Soil water content	θ	0.35	$\text{L}\ \text{L}^{-1}$
Soil dry density	ρ_{dry}	1.6 $(\rho_{\text{wet}} - \theta)$	$\text{kg}_{\text{dw}}\ \text{L}^{-1}$
Mass of soil	M_S	1,000	kg_{ww}
Roots			
Water content of roots	W_R	0.89	$\text{L}\ \text{kg}^{-1}$
Lipid content of roots	L_R	0.025	$\text{kg}\ \text{kg}_{\text{ww}}{}^{-1}$
Transpiration stream	Q	1	$\text{L}\ \text{d}^{-1}$
Root mass	M_R	1	kg_{ww}
1st order growth rate	k_R	0.1	d^{-1}
Leaves			
Shoot mass	M_L	1	kg_{ww}
Leaf area	A_L	5	m^2
Shoot density	ρ_L	1,000	$\text{kg}_{\text{ww}}\ \text{m}^{-3}$
Lipid content leaves	L_L	0.02	$\text{kg}\ \text{kg}_{\text{ww}}{}^{-1}$
Water content leaves	W_L	0.8	$\text{L}\ \text{kg}^{-1}$
Conductance leaves	g_L	86.4	$\text{m}\ \text{d}^{-1}$
Deposition velocity from air	v_{dep}	86.4	$\text{m}\ \text{d}^{-1}$
Growth rate leaves	k_L	0.035	d^{-1}
Transfer with attached soil	R_L	0.01	$\text{kg}\ \text{kg}^{-1}$
Time to harvest	t_L	60	d

concentration vector ***C(0)*** for the next period. This refers to concentrations in any compartment (i.e. soil, roots and leaves).

9.4.5 Input Data for the Root and Leaf Model

Input data are the same for the steady state and the dynamic model version and are mostly taken from the carrot model (Trapp 2002) and the leafy vegetables model (Trapp and Matthies 1995) (Table 9.1).

9.5 Influence of Contaminant-Specific Parameters

The uptake of contaminants into plants and their accumulation depends on contaminant-specific parameters. The importance of some of these parameters is investigated in this section by using the models described before. Also, experimental results are reported, not only with the intention to give a review, but also in order to confirm (or to falsify) the processes and the principal outcomes of the model simulations. Many experimental studies on plant uptake of organic contaminants are available, but most of them are from laboratory or greenhouse experiments, and the results may not always be applicable to field conditions. Results from field studies, however, are less often published. This may be due to high expenses, analytical difficulties or other research priorities. But a reason may also be that the results of uptake studies from outdoors often show a very large variation and are difficult to interpret.

9.5.1 K_{OW} on Accumulation in Roots and Potatoes

Figure 9.5 shows the calculated concentration in roots (steady-state solution Eq. 9.22) and potatoes (Trapp et al. 2007a) for a constant soil concentration of 1 mg kg^{-1} (wet weight). The most relevant contaminant-specific parameter in the root and potato model is the log K_{OW}, which is varied from 0 (polar contaminants) to 8 (super-lipophilic contaminants). The concentrations are compared to the concentrations resulting from the equilibrium partition coefficient approach (*RCF*, Eqs. 9.5 and 9.10) and to the Travis and Arms-regression ("T&A", Eq. 9.3). For very polar contaminants, the concentration in roots and potatoes is predicted to be higher than the concentration in soil. This is because very polar contaminants are mainly found in aqueous phases, and roots and potatoes contain usually more water (up to 95%) than soil (about 30%). For more lipophilic contaminants, the *RCF* (here related to soil, i.e. RCF/K_{SW}) approaches a value of 1, due to the similar sorption capacity of roots and soil.

For polar contaminants, the regression of T&A and the two dynamic models give results close to the equilibrium (*RCF*). With increasing lipophilicity, the predicted concentration decreases and the deviation from equilibrium increases. The

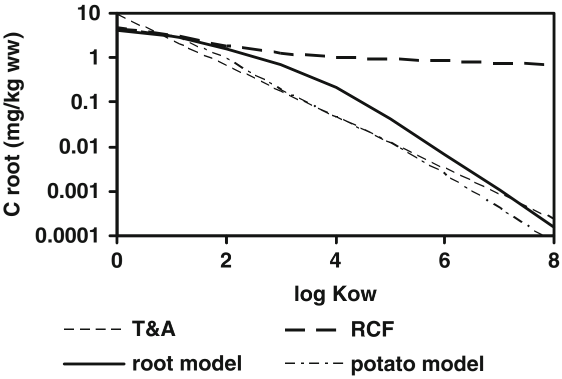


Fig. 9.5 Concentration in root (fresh weight) with varying log K_{OW}, predicted with the regression of Travis and Arms (T&A), the root concentration factor (RCF), the root model and the potato model. Concentration in soil is equal to 1 mg kg^{-1} (wet weight)

root and the potato model include growth dilution, and the effect of this growth dilution increases with increasing log K_{OW}. The *BCF* predicted by the root model at log K_{OW} 7 is more than a factor 1000 below chemical equilibrium. Chemical equilibrium may be found in the peel or close to the peel. But the inner substance of roots and potatoes will have much lower concentrations (Trapp 2002, 2007).

Interestingly, the T&A regression was originally established for above-ground crops. But Fig. 9.5 shows that the predicted bioconcentration in plants is very close to that predicted by the root model (advective uptake) and the potato model (diffusive uptake).

In current chemical Risk Assessment (EC 2003) and in some Risk Assessment tools for contaminated soils, predictions of concentrations in root vegetables are based on the equilibrium approach. This will lead to an overestimation of the concentration in roots and, hence, of human exposure (Legind and Trapp 2009). Recent studies showed that uptake into root vegetables (radishes) may also be from air (Mikes et al. 2009). This has not yet been considered in any available root uptake model.

9.5.2 K_{OW} and K_{AW} on Accumulation of Contaminants in Leaves

Simulations were done with the leaf model described in Section 9.4.3 (Eq. 9.23). A value of 0.1% attached soil was chosen as default. This gives a minimum *BCF* of 0.001 kg kg^{-1} (wet weight) for all contaminants.

Figure 9.6a shows the calculated concentration in leaves for a concentration in soil of 1 mg kg^{-1}, and in air of 0 mg m^{-3}. Log K_{OW} (*x*-axis) is varied for volatile (K_{AW}: 0.2 L L^{-1}), semi-volatile (K_{AW}: 0.001 L L^{-1}) and non-volatile contaminants (K_{AW}: is 10^{-6} L L^{-1}). For all volatile contaminants, the concentration in leaves is low (0.001 mg kg^{-1}) over the whole range of log K_{OW}. This means that their accumulation due to translocation to leaves is low; the plotted concentration is due to attachment of soil particles only (default 0.1%).

The semi-volatile contaminants show some accumulation, up to a *BCF* of 0.01 kg kg^{-1} for the most polar contaminants. With increasing lipophilicity of the

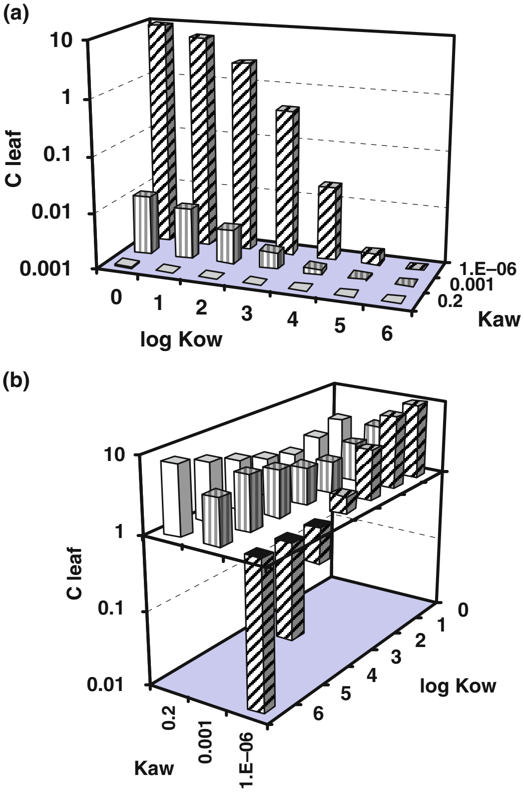


Fig. 9.6 Simulated concentration in leaves for compounds with varying log K_{OW} and varying partition coefficient air-water (K_{AW}), for a concentration in soil equal to 1 mg kg^{-1} (wet weight). Concentration in air is (**a**) 0 and (**b**) in phase equilibrium with soil

contaminants the accumulation decreases. The largest accumulation is seen for non-volatile contaminants, in particular for the polar ones (low K_{OW}). These contaminants are readily soluble in soil pore water, taken up by plants with the soil pore water and translocated to the leaves. From the leaves, the water evaporates, but the contaminants remain. For non-volatile chemicals ($K_{AW} \to 0$), calculated concentrations of polar contaminants in leaves are more than 100 times higher than in soil.

Sulfolane (tetrahydrothiophene 1,1-dioxide) is an organo-sulfur contaminant used as solvent that accumulates in leaves. Sulfolane is neutral, polar (log K_{OW} of –0.77) and non-volatile (K_{AW} of 2.14×10^{-4}). The uptake of sulfolane into wetland vegetation was measured in field and greenhouse studies (Doucette et al. 2005). The measured *RCF* was between 0.3 and 1.4 L kg^{-1} (related to the initial concentration in solution). High translocation to leaves was found. The *BCF* values for shoots were up to 160. Another experiment with sulfolane was done for apples (Chard et al. 2006). The concentration ratio of fruit to soil was 2.8. For leaves, a *BCF* of 652 was found. As far as the authors are aware, this is the highest *BCF* plant to soil that was ever measured, and confirms model predictions in which polar, non-volatile contaminants are best translocated to and accumulated in leaves. Also measured

*BCF*s for hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) (McKone and Maddalena 2007) confirm model predictions for polar non-volatile contaminants.

Contrarily, trichloroethene (TCE) is a volatile chlorinated solvent (K_{AW}: 0.5 L L^{-1}) that does not accumulate in leaves. A study on trichloroethylene uptake by apple and peach trees and transfer to fruit was performed by Chard et al. (2006). No TCE could be detected in fruits, but ^{14}C from unidentified metabolites was found. In leaves, the metabolites dichloroacetic acid (DCAA) and trichloroacetic acid (TCAA) could be detected. The article cites a field study where TCE could be detected in several fruits, but only in traces. Overall, the findings confirm the model prediction (Fig. 9.6a) that volatile contaminants do not show high accumulation in above-ground plant parts like leaves.

9.5.3 Uptake from Air Versus Uptake from Soil

A frequent experimental result is that contaminants are found in moderate or even high concentrations in plants even though concentrations in soil are low (Delschen et al. 1996, 1999; Mikes et al. 2009). This is typically the case when uptake is mainly from air (compare Section 9.3.1). The simulations displayed in Fig. 9.6b were done for identical conditions as for Fig. 9.6a, except that the concentration in air was set to phase equilibrium to soil (i.e., $C_{\mathrm{Air}} = K_{AW} \times C_{\mathrm{Soil}}/K_{SW}$), with concentration in soil equal to 1 mg kg^{-1}). The development of the concentration in plants is completely different from Fig. 9.6a (note that the figure was rotated and the z-axis crosses now at C_{Leaf} equal to 1 mg kg^{-1}). The concentration in leaves is higher than in Fig. 9.6a where there was no contaminant present in the air, in particular for volatile contaminants (K_{AW}: 0.2 L L^{-1}). Also, the concentration is less variable, with most values between 1 and 10 mg kg^{-1}. This is because for most contaminants the system is close to equilibrium in regard to soil with air and air with leaves. An exception are the non-volatile contaminants, their predicted concentration does not change substantially. For the polar and non-volatile contaminants, the calculated concentration in leaves is particularly high.

From Fig. 9.6a and b it can be seen that the partition coefficient between air and water (also known as the dimensionless Henry's Law constant) is a very important parameter for calculation of the accumulation in leaves, because K_{OA} (the ratio of K_{OW} and K_{AW}) determines partitioning into leaves. Leaves have a very high exchange with air (that is their role in plant physiology), and any *volatile* contaminant (with high K_{AW}) will escape from leaves into air and will not accumulate.

The pattern of uptake of contaminants from soil into fruits is very similar (not shown), although the level of concentrations is typically about a factor of 10 lower (Trapp 2007). This means that also, in fruits, polar and non-volatile contaminants have the highest potential for accumulation from soil.

Uptake into fruits of lipophilic contaminants is preferably from air. An example is the transfer of PCDD/F from contaminated sites into field crops, which has been intensively studied. Müller et al. (1994) found an increase of PCDD/F concentrations in the peel of carrots, when grown at a PCDD/F-contaminated site. No increase of the concentration in lettuce and peas was found. In a similar study, it could be shown that the transfer of PCDD/F into apples and pears is exclusively from air (Müller et al. 1993). The results from these studies confirm the model predictions in which uptake of lipophilic contaminants into above-ground plant parts is primarily from air. However, there are exceptions. Hülster et al. (1994) could proof a transfer of PCCD/F into zucchini and pumpkins in field experiments, which was much higher than for other fruits.

Another example is benzo(a)pyrene (BaP). The range of measured concentrations of BaP in food crops is quite large, from 0.01 to 48 μg kg^{-1}. Generally, leafy vegetables and lettuce had the highest concentrations, followed by grains, potato and root vegetables (Kazerouni et al. 2001; Samsøe-Petersen et al. 2002; SCF 2002). This indicates air as the primary source for contamination of vegetables with PAHs.

Uptake of organochlorine pesticides and polychlorinated biphenyls (PCBs) from soil and air into radishes were measured at a contaminated field site in the Czech Republic (Mikes et al. 2009). Root concentration factors (*RCF*) and bioconcentration factors for leaves (BCF_L) were determined by linear regression. The transfer from soil into leaves (average BCF_L 0.08 kg $\text{kg}_{\text{dw}}{}^{-1}$) was rather constant for all contaminants and could best be explained by soil particle attachment. Generally, uptake from air was higher than uptake from soil. Uptake from air into radish roots and bulbs was observed. The example in Section 9.3.1 was taken from this study.

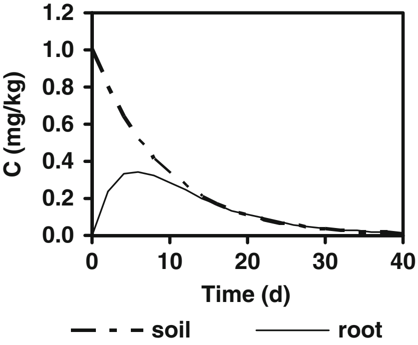
9.5.4 Dissipation from Soil

The simulations above were done for the steady-state situation, with constant concentration in soil. However, contaminants can rapidly dissipate from soil by several removal processes, such as degradation, plant uptake, leaching, volatilization and sequestration.

An example is shown in Fig. 9.7. Ortho-xylene is a moderately lipophilic contaminant which is rapidly degraded in soil. Søvik et al. (2002) give a first-order degradation rate constant of 0.11 d^{-1}, i.e. a half-life of about 1 week. After a few weeks, the concentration in soil has approached zero. The simulated concentration in roots closely follows this pattern: About four days after the pulse input, the calculated concentration in root is already close to equilibrium to soil, and parallel with the concentration in soil, concentration in roots falls to very low values within a few weeks, in any case before the harvest. A steady state simulation with constant soil concentration would overpredict concentrations in the harvest product by several orders of magnitude.

The degradation or total dissipation rate is a key variable. To some extent, elimination in soil is related to physico-chemical properties. Bacterial degradation depends on the bioavailable fraction of the contaminants (Alexander 2000; Reichenberg and Mayer 2006), which is typically higher at low sorption and low sequestration. Plant uptake and leaching are important for soluble contaminants

Fig. 9.7 Simulated concentration of o-xylene in soil and roots as a function of time, with a degradation rate of 0.11 d^{-1}



with low K_{OW}. Volatilization to air depends on K_{AW} (or K_{OA}). Therefore, polar and volatile contaminants (low K_{OW}, high K_{AW}) show shorter residence time in soil. The actual dissipation rate, however, cannot be predicted from physico-chemical properties alone but is an input data usually derived from experiments.

9.5.5 Impact of pKa and pH on Uptake of Ionisable Contaminants

Ionisable contaminants, i.e. acids, bases, zwitterions or amphoters, may be present in soils as neutral or ionic molecules. The neutral and the ionic molecules have completely different physico-chemical properties. The ion is usually much more polar and water soluble and has a very low vapor pressure compared to the neutral species. Thus, the ion is a polar and non-volatile contaminant, and as such has a high potential for accumulation in plants. On the other hand, cations have a strong tendency to adsorb to soil organic matter and/or clay (Franco and Trapp 2008). Besides, charged contaminants cross biomembranes slowly (Trapp 2004), which limits their uptake.

A process that may lead to high accumulation of ionisable contaminants is the ion trap. This principle is described in Fig. 9.8, for an acid. If the pH outside in the soil pore water is low, a weak acid is at least partly neutral. The uptake into the cell is rapid (provided that the contaminant is not too hydrophilic). The pH in cell sap (cytosol) is about 7 to 7.5, in xylem about 5.5, and in phloem about 8. In particular in cytosol and phloem, weak acids dissociate and form the anion. The membrane permeability of the anion is very low, and reverse diffusion is slow. As a consequence, the anion is trapped inside the cytosol or the phloem. The same process occurs for a base, when the cell sap is acid, compared to the outside soil pore water, for example in the vacuoles of plant cells and in the xylem (pH is 5.5).

Summarizing, the ion trap requires a gradient in pH between outside and inside of the plant, so that the outside pH is lower (for bases: higher) than the inside pH. Second, a pKa at or somewhat below (for bases: above) the outside pH is necessary, so that a relevant fraction of contaminants is neutral outside, but most of it is

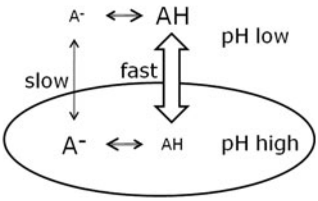


Fig. 9.8 Principle of the ion trap. AH is the neutral contaminant (here an acid), and A^- is the ion (here an anion)

ionic inside. This means that weak acids (pKa 2–6) may accumulate in the alkaline phloem (and are in consequence transported to fruits), while weak bases (pKa 6–10) tend to accumulate in the acidic xylem (and are translocated to leaves). Examples of such contaminants are herbicides like 2,4-D (Shone and Wood 1974) and sulfonylureas (Fahl et al. 1995), and drugs like the base trimethropin (Boxall et al. 2006) and the acid ibuprofen.

Briggs et al. (1987) determined the *RCF* and *TSCF* of weak organic acids with log K_{OW} between 0.06 and 4.51 and pKa-values near 3. At pH 7 in external solution, *RCF*-values were low, between 0.5 and 4.5 L kg^{-1}. At pH 4, the *RCF*-values were higher and ranged from 2.6 to 72 L kg^{-1}. Similar, the *TSCF*-values of weak acids at pH 7 were low, 0.04 to 0.05 L L^{-1}, while at pH 4 *TSCF*-values up to 4.2 L L^{-1} were found. Briggs et al. used the ion trap process to interpret their results. Chamberlain et al. (1998) studied the uptake into roots and translocation to shoots of two weak bases, i.e. dodemorph (pKa 7.8) and tridemorph (pKa 7.4). At pH 5, *RCF*-values of dodemorph were < 10 L kg^{-1}, and of tridemorph about 20 L kg^{-1}. With increasing pH, *RCF* increased to 49 L kg^{-1} for dodemorph and 183 L kg^{-1} for tridemorph at pH 8. In parallel, the *TSCF* increased from < 1 L L^{-1} for both bases at pH 5 to 24 L L^{-1} for dodemorph and slightly below 10 L L^{-1} for tridemorph at pH 8. To our knowledge, the *TSCF* of dodemorph in this experiment was the highest *TSCF* ever determined. A similar study was done by Inoue et al. (1998), who also found a large increase of uptake and translocation of weak bases in plants with increasing pH. The results from these studies underline the importance of pH in the soil pore water for uptake and translocation of weak electrolytes. However, the results have not yet been confirmed by field studies.

9.6 Influence of Plant-Specific Parameters

As pointed out in the introduction, there are a very high number of plant-contaminant combinations, and the simulations done for a generic plant are not necessarily valid for the whole plant empire and all contaminants. In this section, some properties of plants are discussed, which vary and may lead to different accumulation of contaminants.

9.6.1 Crop Types and Uptake Pathways

The crop type has a very large impact on uptake processes, e.g. roots and potatoes are in close contact to soil, while apples are not. As a consequence, the accumulation of contaminants from soil is much lower in apples and other tree fruits, whereas the accumulation by uptake directly from air is higher for fruits. It is obvious that crop-specific models will give more realistic predictions of concentrations.

Even between different species of the same crop type differences in uptake can exist. For zucchini and pumpkin (*Cucurbita pepo*), both members of the plant family *Cucurbitaceae* and the genus *Cucurbita*, root uptake and subsequent translocation to shoots and fruits was the main uptake route of PCDD/F, probably due to root exudates which mobilize lipophilic contaminants (Hülster and Marschner 1995) and an increased mobility in the transpiration stream (Gent et al. 2007). However, fruits and leaves from other plant species, even for the closely related cucumber plant (*Cucumis sativus*), were mainly contaminated by airborne PCDD/F (Hülster and Marschner 1994; Hülster et al. 1994). The exact reason for the ability of cucurbita plants to extract lipophilic contaminants from soil and transport them with the xylem sap is yet unknown, but it was observed repeatedly, and for a number of lipophilic organic contaminants, such as p,p′-DDE (White 2002), DDT (Lunney et al. 2004), PCB (Whitfield Åslund et al. 2007) and PBDE (Mueller et al. 2006). For phenanthrene, unusual high adsorption to a range of plant species was observed (Zhu et al. 2007).

9.6.2 Physiological Parameters

The importance of physiological plant-specific parameters, such as transpiration rate, leaf area, conductance, water and lipid contents as well as growth rate, depends largely on the properties of the contaminant. For water soluble contaminants, which are rapidly translocated from soil to leaves, the transpiration rate is among the most important parameters, since the accumulation in leaves is almost directly dependent on the transpiration.

This is illustrated with the example of carbofuran in Fig. 9.9. The log K_{OW} of this contaminant is 1.6–2.07 (Trapp and Pussemier 1991). Two calculations were done, one with the standard transpiration of 1 L d^{-1}, the other with an increased transpiration of 5 L d^{-1}. The simulated concentration in roots remains nearly constant with increasing transpiration, only the time period until steady state is reached is shorter. However, the breakthrough of contaminant into leaves is faster and the calculated concentration in leaves is a factor of 5 higher for the increased transpiration rate due to their direct relationship. The translocation of carbofuran in bean plants was determined experimentally, and the concentration pattern as well as the close relation between transpiration and concentration in leaves was confirmed (Trapp and Pussemier 1991).

In turn, the amount of transpired water depends on temperature, humidity, leaf area and stomata permeability. However, the transpiration depends also on the

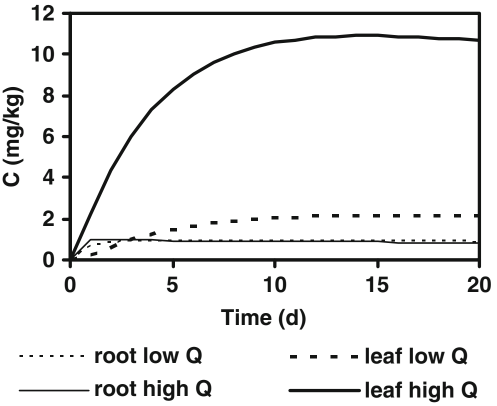


Fig. 9.9 Simulation of the concentration of carbofuran in plants as a function of time, with low transpiration (Q is 1 L d^{-1}) and high transpiration (Q is 5 L d^{-1}). Dynamic model version, data Table 9.1

availability of water. Students in a field course at the Technical University of Denmark noticed that concentrations of chlorinated solvents (PCE and TCE) measured in wood cores taken from trees growing at the Glostrup site (Denmark) were much lower in the very dry June 2008 than in the very wet June 2007 (Mette Broholm, personal communication). A plausible explanation is a reduced transpiration due to water stress in 2008. Transpiration in field settings can be estimated using the Penman-Monteith model by Penman (1948) and Monteith (1964) (cited in (Monteith 1995)).

Leaves grow and contaminants entering the leaf with the transpiration stream are not uniformly distributed in the leaf, but tend to accumulate in the leaf tip (Doucette et al. 2005).

9.6.3 Plant Morphology and Collection Efficiency for Particles

The resuspension of contaminated soil particles from soil to plant surfaces is an important transport mechanism for lipophilic contaminants. This transfer pathway was well studied using insoluble radionuclides, and large variations were found in soil attachment among plant species. Soil particles may reach leaf surfaces mainly by three mechanisms, namely rain splash, wind erosion and erosion due to agricultural practice (ploughing, harvesting et cetera). Table 9.2 shows measured values for the transfer of radionuclides (Li et al. 1994). The transfer range varies from 1.1 to 260 mg soil $g_{\mathrm{plant}}{}^{-1}$ (dry weight), depending on the species. But also, the intraspecies variation was considerable, as can be seen from the standard deviations given in Table 9.2. The highest value was found for lettuce, with 260 mg soil $g_{\mathrm{plant}}{}^{-1}$ (dry weight), probably due to the small size of the plant (< 40 cm), the leaf morphology where leaves are collecting particles to the interior of the plant, and the surface structure. It was also demonstrated that particle-bound organic contaminants migrate from the leaf surface to the interior of the leaf (i.e. are overgrown by surface waxes), from where they can not be washed off (Kaupp 1996). A BCF for radish leaves to soil of 0.08 kg $kg_{\mathrm{dw}}{}^{-1}$ was recently measured at a Czech site for

Table 9.2 Transfer of attached soil particles to leaf surfaces (Li et al. 1994); expressed as mg soil $g_{\text{plant dw}}{}^{-1}$ ± standard deviation

Plant species	Amount of soil (mg soil $g_{\text{plant dw}}{}^{-1}$)
Lettuce	260 ± 100
Turnip	32 ± 11
Grass	18 ± 48
Tomato	17
Broccoli	10 ± 8.1
Bean leaves	9.5
Grass (Lolium)	5.8
Wheat	4.8
Sun flowers	2.6 ± 0.9
Tobacco	2.1 ± 0.6
Soy bean	2.1
Maize	1.4
Cabbage	1.1 ± 1.1
Average	28.2
Average (for fresh weight)*	1.4

*Based on 95% water content of plants

persistent lipophilic contaminants and was contributed to soil particle attachment (Mikes et al. 2009).

A default value of 10 mg soil $g_{\text{plant}}{}^{-1}$ (fresh weight) for transfer of attached soil was chosen for lettuce in the New Model Framework for dietary exposure of children and adults (Legind and Trapp 2009) (for lettuce, $\rho_{L,wet} = 1\ \text{kg}_{\text{fw}}\ \text{L}^{-1}$ and $\rho_{L,dry} = 0.2\ \text{kg}_{\text{dw}}\ \text{L}^{-1}$, so this corresponds to 50 mg soil $g_{\text{plant}}{}^{-1}$ (dry weight)). For grains, 1 mg soil $g_{\text{plant}}{}^{-1}$ (fresh weight) transfer with attached soil particles was assumed, due to pollution at harvest. In models predicting the exposure to radionuclides, the default value for the transfer of attached soil is 5 mg soil $g_{\text{plant}}{}^{-1}$ (fresh weight) (25 mg soil $g_{\text{plant}}{}^{-1}$ (dry weight)) for pasture grass, and 1 mg soil $g_{\text{plant}}{}^{-1}$ (fresh weight) for other plants (Paretzke and Garland 1990).

In systematic experiments, Delschen et al. (1996, 1999) investigated the uptake and the uptake pathways of PAHs (polycyclic aromatic hydrocarbons) into vegetables and crops. They found that uptake of PAHs can both be from contaminated sites and from the atmosphere. In some experiments, the soil was covered with fine sand or a gas-permeable foliage. This avoided soil particle attachment, but allowed for volatilization with subsequent adsorption to leaves. In fact, this coverage reduced the uptake of 15 PAHs from soil almost completely, indicating soil particle attachment to leaves as a major transfer pathway of PAHs (in particular benzo(a)pyrene and dibenz(a,h)-anthracene) from soil to vegetables. Thus, crops with harvested parts close to the soil surface have the highest affinity for accumulation, because soil particle attachment by rain splashing rarely affects plants with a height above 40 cm (Dreicer et al. 1984; Li et al. 1994). No systemic uptake of PAHs via plant roots was found. Concentrations of PAHs in peels of potato and carrots from organic farming were generally higher than in the core of potatoes and carrots (Zohair et al. 2006). The *BCF* for individual PAHs ranged from 0.0002 to 0.3 kg kg^{-1} and decreased with log K_{OW} (Trapp et al. 2007a; Zohair et al. 2006). The results from the experiments are in good accordance with the model predictions.

Note that the ingestion of soil attached to the daily vegetable diet may be higher than the current estimate for direct soil ingestion used in Exposure Assessments for humans. The direct ingestion of soil is 50 mg/day for adults and 100 mg/d for children (ECETOC 2001). The average lettuce consumption in Denmark is 6 g/d for children and 9 g/d for an adult woman. A transfer of 1% (10 mg soil $g_{plant}{}^{-1}$ (fresh weight)) due to attached soil corresponds to 60 mg/d and 90 mg/d soil ingestion with lettuce only. Other vegetables and fruits are consumed at 389 g/d (children) and 475 g/d (mother). If on average 0.1% soil were attached, this corresponds to 389 mg/d and 475 mg/d of soil, which is much more than the value of direct ingestion of soil and deserves consideration in Human Health Exposure Assessment.

9.6.4 Variation of Partition Coefficients

The *RCF* regressions for barley (Briggs et al. 1982) and bean plants (Trapp and Pussemier 1991) differ (Fig. 9.3). Also, the partition coefficients between leaves and air, K_{LA}, vary with plant species. The K_{LA} depends both on plant-specific parameters (such as plant lipid content and empirical exponent b (Eq. 9.10)) and contaminant-specific parameters (K_{OW} and K_{AW} or K_{OA}). Kömp and McLachlan (1997) found differences in the uptake of PCB between five different plant species of up to a factor of 20. There was a linear relationship between log K_{PA} and log K_{OA} values within each plant species, but the slopes of the regressions ranged from 0.57 to 1.15. The standard value for leaves (slope $= b = 0.95$, Eq. 9.9) was determined by Briggs et al. (1983) and lies in this range.

9.6.5 Permeability

The velocity of diffusive uptake from air and loss to air is controlled by the permeability of leaves (synonyms exchange velocity, conductance or transfer velocity). Three major resistances control the exchange: air boundary layer resistance, stomata resistance and cuticle resistance (Riederer 1995). Their importance depends on the chemical properties but varies for each leaf. Volatile and gaseous contaminants will prefer stomata for the entry to or escape from plants, while lipophilic contaminants will preferably diffuse through cuticles. The stomata resistance for water and contaminants can be calculated from the transpiration rate, leaf area, temperature and humidity (Trapp 1995, 2007). The cuticle resistance depends on thickness and diffusivity inside the cuticle (Schönherr and Riederer 1989). The latter is highly variable with species and temperature (Buchholz et al. 1998): at high temperatures, the surface waxes liquidize and get more permeable.

A method for estimation of conductance g (m s^{-1}, related to gas phase) is given by Trapp (1995, 2007). Values for g estimated with this method range from 9×10^{-3}

m s^{-1} for non-volatile, lipophilic contaminants to 0.2×10^{-3} m s^{-1} for volatile or polar contaminants. A default value of 10^{-3} m s^{-1} was suggested, to avoid lenghty calculations (Trapp and Matthies 1995).

9.6.6 Particle Deposition

A similar variability can be observed for the particle deposition rate. Particle deposition depends on the fraction of contaminant adsorbed to particles, the deposition velocity of particles, as well as characteristics of the leaves. The fraction adsorbed to particles, f_P, is often calculated from the *Junge* equation, using vapor pressure as input data (Junge 1977; cited in EC 2003). Newer methods are based on the K_{OA} (Cousins and Mackay 2001).

The particle deposition rate for wet and dry deposition, v_{dep}, changes with the diameter of the particles. Default values are given in the German TA-Luft and are listed in Table 9.3. Generally, diameters and thus deposition velocities are higher close to the source of emission. At larger distance the coarse particles have settled, and finer particles remain. The deposition rate of fine particles is similar to the conductance of gases, and the default value of v_{dep} (10^{-3} m s^{-1}) is the same as for conductance g. Thus, uncertainties in f_P and v_{dep} are not crucial for the calculated result.

Bakker et al. (1999) studied the deposition of polycyclic aromatic hydrocarbons (PAHs) on leaves from three closely related *Plantago* species and found variations up to a factor of five, which could be explained solely by plant morphology and surface structure.

Table 9.3 Deposition velocities of particles (TA-Luft 1986)

Diameter (μm)	v_{dep} (mm s^{-1})
< 5	1
5 to 10	10
10 to 50	50
> 50	100
unknown	70

9.6.7 Metabolism in Plants

Rapid metabolism in plants will significantly decrease the contaminant concentration in plants. Transformation of contaminants may occur in the rhizosphere, inside the plant and on the leaf surface. Often, it is difficult to differentiate between metabolism by fungi or bacteria living on and in plants and metabolism by plant cells. Stimulation of bacterial degradation in the root zone has been observed frequently and is a key process in phytoremediation projects (Trapp and Karlson 2001). Therefore, it may happen that non-persistent contaminants do not reach the roots but are degraded by bacteria living on or in the vicinity of roots.

The role of the plant enzyme system is detoxification, and there is an upper limit, beyond which plants suffer from toxic effects and probably die (Trapp et al. 2007b). This contamination of plants may limit activities in gardening, agriculture and forestry on contaminated sites.

Metabolism by plants has been described by the green liver concept, because plant metabolism rather resembles the processes in the animal liver than the bacterial metabolism (Sandermann 1994). The first step (phase I reaction) is typically an oxygenation with cytochrome P-450, followed by conjugation reactions (phase II) with glutathione-S-transferases (GST) (Barret 1995; Pflugmacher and Schröder 1995). Unlike animals, plants are not able to excrete conjugates via the urine. Instead, phase III of plant xenobiotic metabolism involves storage of soluble conjugates in the vacuole and of insoluble conjugates in the cell wall (Komossa et al. 1995). This may lead to so-called bound residues. These bound unextractable residues resist solubilization in common laboratory solvents and are therefore not accessible to standard residue analysis. It was found that bound residues can be present in larger amounts than the parent contaminant and could therefore represent a source of significant consumer exposure (Sandermann 2004).

Little is known about metabolism rates of contaminants by vegetation. Cyanide (HCN) was used as model contaminant to study the variation of rates among plant species. Even though inorganic, cyanide behaves like an organic contaminant in terms of lipid solubility, volatility and metabolism. The removal of free cyanide followed Michael-Menten kinetics (Larsen et al. 2004). Adding Michaelis-Menten kinetics to the mass balance equation for roots (Eq. 9.18) leads to the following non-linear equation:

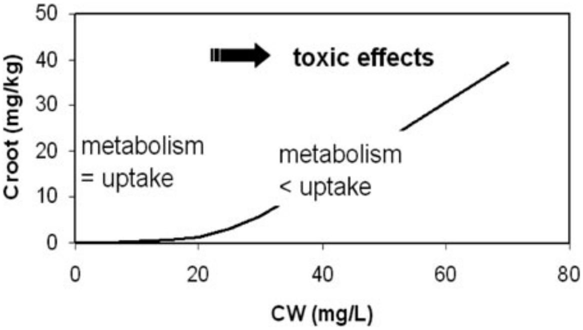
$$\frac{dC_R}{dt} = \frac{Q}{M_R} \times C_W - \left(\frac{Q}{M_R \times K_{RW}} + k_R\right) \times C_R - \frac{v_{\max} \times C_R}{K_M \times K_{RW} + C_R} \qquad (9.38)$$

where v_{max} (mg kg^{-1} d^{-1}) is the maximal metabolism velocity of the contaminant and K_M (mg L^{-1}) is the half-saturation constant and (Larsen et al. 2005).

At a low external concentration in soil pore water, all contaminants that are taken up are metabolized (Fig. 9.10). At higher concentrations, however, the enzyme system is overloaded. Then, uptake is linearly related to the external concentration. This was shown experimentally for free cyanide (HCN) by Larsen et al. (2005). A non-linear *BCF* relation indicating enzymatic activity of plants was found repeatedly, e.g. for phenol (Ucisik and Trapp 2006) and salt, NaCl (Trapp et al. 2008). In these cases, the *BCF* was low at a low external concentration in soil pore water, but increased at higher external concentrations.

The Michaelis-Menten parameters K_M and $v_{\max}$ varied with plant species, but less than expected. Values of $v_{\max}$ of 12 species from nine plant families were found in a relatively narrow range between 6.7 and 21.9 mg CN $\text{kg}_{\text{plant}}{}^{-1}$ h^{-1} and were normally distributed with a mean of 13 mg CN $\text{kg}_{\text{plant}}{}^{-1}$ h^{-1} (Yu et al. 2004). The authors concluded that the variation of metabolism rates between plant species was

Fig. 9.10 Calculation of concentration in roots of cyanide as a function of the concentration in external aqueous solution (C_W), using Michaelis-Menten kinetics for metabolism; v_{max} is 6.9 mg kg^{-1} h^{-1}; K_M is 0.44 mg L^{-1} (Larsen et al. 2005)



small, and that it is likely that all investigated plants used the same enzyme for detoxification. Another sign for common pathways and rates of metabolism in plant species is the difficulty to find selectively acting herbicides: it is rare that a toxin hits weeds but not crops (Börner 1995).

9.7 Environmental Variables

Which contaminants are taken up into plants and the variation in the affinity for uptake, and hence accumulation, depends also on environmental parameters. The importance of some of these parameters is investigated in this section by using the models described before and some experimental results.

9.7.1 Climate

Principally, the models can be adapted to all climatic conditions at which plants grow. In particular temperature influences uptake and fate of xenobiotics in various ways. At higher temperatures plant physiological processes such as transpiration, growth and metabolism are stimulated. According to *Claudius-Clapeyron*'s law, the vapor pressure increases exponentially with temperature, so the vapor pressure of water in air is 872 Pa at 5°C but 5635 Pa at 35°C, i.e. a factor of 6.5 higher. The potential transpiration is directly proportional to the vapor pressure of water. Similarly, metabolic reactions are stimulated at higher temperatures. For example, the removal rate of cyanide was 3–5 times higher at 30°C than at 11°C (Yu et al. 2005a, 2007). Also, contaminant properties change with temperature. Like the vapor pressure, the partition coefficient air to water, K_{AW} also increases exponentially with temperature. For chlorinated solvents, such as TCE, the K_{AW} at 5°C is about a factor of 3 below that at 35°C (Kühne et al. 2005). This means less accumulation in leaves and more volatilization to air at higher temperatures.

9.7.2 Bioavailability

The concentration in the soil pore water is for neutral organic contaminants calculated from the K_{OC}. The bioavailability of contaminants may be reduced due to aging (Alexander 2000). At contaminated sites the soil pore water concentrations can be much lower and sorption coefficients can be much higher than equilibrium partitioning models predict (Ter Laak et al. 2006). Cations are attracted by the electrical potential of living cells, but also adsorb to soil organic carbon and to negatively charged clay particles (Franco and Trapp 2008). This reduces their bioavailability and, hence, uptake. Also, a depletion of contaminants in soil due to uptake into plants should be taken into account. In Chapter 16 an extended description of bioavailability is given.

9.7.3 Soil pH

Soil pH directly affects the speciation of acids and bases, as described by the *Henderson-Hasselbalch* law ($\log [A^-]/[HA] = pH - pKa$). Uptake of anions is generally lower than of neutral molecules, due to electrical repulsion and slow transfer across membranes. Both weak acids and bases can undergo the ion trap process (Section 9.5.5). This will lead to an accumulation of weak acids from acidic soils, and of bases from alkaline soils. This conclusion is based on a model prediction and has not yet been confirmed by experiments. The pH also has an indirect effect on uptake: many plants do not grow well outside their optimum pH range. Extreme pH (high or low), will lead to reduced growth, and this may be accompanied by reduced uptake of contaminants.

9.7.4 Uncertainties in Predictions

Sections 9.5, 9.6 and 9.7 list parameters and variables that influence the uptake of contaminants into plants. Most likely, this list is far from complete. This may explain why under some conditions, in some experiments, a high uptake of a contaminant into a plant may be found, while this may not be the case in the next investigation, under other conditions. It also explains why data from experiments with plants often suffer from large scatter. Besides, care must be taken when results obtained from uptake studies are translated to other crop types, other climates, and other agricultural practices.

Models may help to design and interpret uptake experiments, in indicating relevant processes and pathways, and hence in translating results to other conditions. But due to the large number of parameters and their high variability in space and time, these models can not be expected to give exact results. Some studies tested the validity of model approaches (Fryer and Collins 2003; Legind and Trapp 2009; McKone and Maddalena 2007; Rikken et al. 2001; Trapp and Schwartz 2000).

Often, the results were ambiguous, because the influence of concentrations in air could not be quantified (compare Section 9.3.1), but also due to large uncertainties in measured as well as predicted concentrations in plants (McKone and Maddalena 2007). The primary role of models is to indicate relevant processes and the potential for accumulation of contaminants in plants. Good decision making needs to consider both model predictions and experimental data.

In a critical review of uncertainties related to soil-to-plant bioconcentration factors by McKone and Maddalena (2007), *BCF*-values for the explosive RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine) are listed. RDX is a quite persistent, relatively polar (log K_{OW}: 0.87) and non-volatile contaminant (K_{AW}: 6.3×10^{-8} atm $\mathrm{m}^3\ \mathrm{mol}^{-1}$). *BCF*-values range from 0.06 (mg $\mathrm{kg_{ww\ plant}}^{-1}$: mg $\mathrm{kg_{dw\ soil}}^{-1}$) in bean pods and 0.07 (mg kg $_{\mathrm{ww\ plant}}{}^{-1}$: mg L water^{-1}) in bean fruits at the lower end to 19.2 (mg $\mathrm{kg_{ww\ plant}}^{-1}$: mg $\mathrm{L_{water}}^{-1}$) in bean leaves and 28.6 (mg $\mathrm{kg_{dw\ plant}}^{-1}$: mg $\mathrm{kg_{dw\ soil}}^{-1}$) in carrot shoots at the upper end. Overall, fruits tend to have lower *BCF* values than roots, and leaves have the highest accumulation, but there is a very large variation in the data. The *BCF* values for fruits range from 0.07 to 5.50, for roots from 0.08 to 4.50 and for leaves from 0.30 to 28.6, i.e. the maximum *BCF* is 79 fold, 56 fold or 93 fold larger than the minimum *BCF*. Even though the units are not uniform (i.e. a mix of wet weight and dry weight, soil and water), this variation is quite impressive for one single contaminant and similar crop type.

9.8 Uptake Potential of Specific Substance Classes

This section gives a short classification of several frequently found soil and groundwater contaminants with respect to their uptake into plants. It is based on the model simulations, measured data and the process review above, but also on the experience of the authors.

9.8.1 Chlorinated Solvents (PCE, TCE and Others)

Chlorinated solvents such as perchloroethene (PCE) and trichloroethene (TCE) are probably the most frequently found groundwater contaminants world-wide. They are water-soluble, relatively persistent and very volatile. Uptake of chlorinated solvents into tree trunks has been reported frequently (Chard et al. 2006; Larsen et al. 2008; Sorek et al. 2008; Vroblesky et al. 1999). However, these contaminants are highly volatile and escape rapidly from branches, small stems and leaves into the air (Baduru et al. 2008). Thus, in fruits, only metabolites of chlorinated solvents could be detected (Chard et al. 2006). Accumulation in root crops and potatoes is likely, though not described yet. During storage and cooking, chlorinated solvents could be eliminated from food stuff, due to their high volatility. Moreover, these contaminants are not expected to be present in high concentrations in top soil, i.e. the root zone, due to volatilization to the air and leaching to the groundwater.

9.8.2 Gasoline Contaminants

Gasoline is a mixture of light petroleum products (for example alkanes such as hexane, and aromates such as benzene and toluene) plus additives like ethanol and MTBE. They frequently occur in the groundwater due to leaching storage tanks. Gasoline contaminants are volatile and usually rapidly degraded. Thus, concentrations in aerated soil are generally low (Fig. 9.7). Accumulation in crops from soil has not been described, except in laboratory experiments. An exception is MTBE, which is less volatile and more persistent. MTBE was not metabolized in any of 24 plant species tested (Trapp et al. 2003). Uptake of benzene from air is rapid but levels are low, due to a low K_{OA}, and insignificant for human exposure (Collins et al. 2000).

9.8.3 Heavy Petroleum Products

The heavy fraction of petroleum products consists of long-chain alkanes and some polycyclic aromatic hydrocarbons. These contaminants are lipophilic, volatile and degradable under aerobic conditions. Significant uptake from soil into plants has not been observed and is not expected from the properties of petroleum components.

9.8.4 Polycyclic Aromatic Hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) is a group of contaminants with two (naphthalene) to seven or more condensed aromatic rings with wide-spread occurrence in the environment from incomplete combustion. Most of these contaminants are very lipophilic (log K_{OW} 5−7) and have a very low water-solubility. Uptake into plants with transpiration water is therefore unlikely. Contamination of plants is mainly via attached soil particles or from air (Delschen et al. 1999). Uptake from soil into the peel of carrots or potatoes may occur. Solubilization by cucurbita species (pumpkin and zucchini) is likely.

9.8.5 Persistent Organic Pollutants POPs

Persistent Organic Pollutants (POPs) are chlorinated contaminants like the pesticides DDT, dieldrin and lindane, but also polychlorinated biphenyls (PCB) and polychlorinated dibenzo-*p*-dioxins and -furans (PCDD/F). POPs are persistent, semi-volatile and lipophilic. Similar to PAHs, the uptake with transpiration water into plants is very unlikely, while attached soil particles can lead to contamination of plants. Solubilization with subsequent translocation by cucurbita species (pumpkin and zucchini) has been observed repeatedly.

9.8.6 Explosives

Explosives such as trinitrotoluene (TNT) and hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) are relatively polar contaminants with low vapor pressure. High uptake into plants has been observed (McKone and Maddalena 2007, Thompson et al. 1998). Inside plants, a reduction of the nitro groups to amino groups is likely and would lead to an increase of the toxic potential. Furthermore, bound residues may be formed.

9.8.7 Phenols

Phenol is the common name for a class of contaminants with a phenolic ring. A variety of chemical groups may be attached to the ring. This leads to contaminants with a large variety of properties (nitrophenols, anilines, chlorinated phenols). Phenols are less volatile than the aromatics benzene, toluene, ethylbenzene and xylene (BTEX). The contaminant phenol as well as monochlorophenol were degraded rapidly in the root zone (Ucisik and Trapp 2006, 2008), while 2,4-dichlorophenol was taken up into roots and translocated to stem, and to some extent also to leaves (Ucisik et al. 2007).

9.8.8 Cyanides

Cyanides are contaminants with a CN-group. They are often found at (abandoned) gasworks sites and gold mines. Free cyanide (HCN, CN^-) is rapidly taken up into roots and metabolized. An accumulation in healthy plants was not observed, because free CN is highly toxic (Larsen et al. 2004, 2005; Yu et al. 2004, 2005b). Iron-complexed cyanide (ferri- and ferrocyanide) is taken up by plants and translocated upwards, while slow metabolism was observed (Ebbs et al. 2003; Larsen and Trapp 2006; Samiotakis and Ebbs 2003). Field measurements at a former gas works site showed no accumulation of total cyanide (sum of free and complexed CN) in leaves (Trapp and Christiansen 2003, and own results, unpublished). Surface contamination of crops by solid iron cyanide (such as Prussian blue) is possible and has been observed (“blue strawberries”).

9.9 Monitoring of Contaminants in Soils and Shallow Aquifers with Vegetation

The uptake of soil and groundwater contaminants into plants has also positive aspects: contaminants such as chlorinated solvents in soils and shallow aquifers can be monitored using plant samples. In principle, every tree could be considered as a well, a pump and a passive sampler, all in one. Plant samples can be taken from stems of trees (Fig. 9.11) and analyzed to identify and map subsurface contamination. The method is rapid and cheap. Screening an area suspect of contamination in the upper aquifer with the tree core method prior to boring observation wells can support the determination of optimal location of these wells, instead of arbitrarily placing wells. This reduces the number of wells necessary for the investigation and thus reduces overall costs. Besides, tree cores can be sampled at sites which are

Fig. 9.11 Demonstration of the tree core method by one of the authors

difficult to access with heavy equipment, for example gardens in residential areas or dense forests. Even better, the effect on trees of the contaminants in a plume is reflected by the size of tree rings, and dendrochronology can be used to determine when a plume passed below the tree (Balouet et al. 2007).

The potential of this monitoring procedure is large. Tree core sampling as monitoring method was successfully applied for chloroethenes (Gopalakrishnan et al. 2007; Vroblesky et al. 1999). Larsen et al. (2008) found a good agreement between the plume delineation by traditional well drilling and by tree core sampling, but only monotone (and not linear) relations. The authors could also use the method to monitor Natural Attenuation of a PCE spill. Sorek et al. (2008) even detected several before unknown spills in Tel Aviv by a random sampling of trees in the city. But there are also limitations. Tree core sampling should rather be used for assessing the presence of contaminants than for determining exact subsurface concentrations. Because from tree core measurements alone, precise evaluation of subsurface contamination will not be possible, due to varying concentrations in wood (Larsen et al. 2008; Sorek et al. 2008).

The method could also be successful for some heavy metals such as cadmium and copper, and it may also work for MTBE, perchlorate and dichlorophenol. For a number of contaminants the method was not successful, to mention are BTEX, PAHs (except naphthalene), free and iron complexed cyanide, long-chain alkanes (generally petroleum products), phenol, monochlorophenol, iron and lead.

A free guide to vegetation sampling for screening of subsurface pollution is available from the web (Trapp et al. 2009).

9.10 Conclusions

What can be concluded from this review? An important conclusion is that the *polar* contaminants are preferably translocated from soil and accumulate in leaves and fruits. Lipophilic contaminants will be retained in soil, roots and the lower stem and will not reach the leaves or fruits in significant amounts. Thus, *polar and non-volatile* contaminants, i.e. contaminants with low K_{OW} and K_{AW}, have the highest accumulation potential from soil into plants. This is in conflict with the usual concept of bioaccumulation, where high bioaccumulation is assumed to occur for highly lipophilic contaminants (EC 2003; Mackay and Fraser 2000). Furthermore, under background conditions, all contaminants except the polar and non-volatile contaminants will preferably be taken up from air and the concentration in soil does not have much impact on the concentration in leaves, unless it is far above chemical equilibrium.

The number of available experimental data on plant uptake of contaminants increases. Often, model concepts and measurements are in good agreement, but we may also encounter surprises. Plants are living organisms, and the uptake processes are complex and variable. Uncertainties in the model predictions are high, and large variations may be expected in predicted, but also in experimental data, due to the variations of contaminant-specific, plant physiological and environmental parameters as described in this chapter.

Acknowledgement We thank our editor, Frank Swartjes, for his initiative, patience and help. Our work is supported by the European Union 6th Framework Program of Research, Thematic Priority 6 (Global change and ecosystems), contract number GOCE-CT-2007–036976, project 2-FUN, and contract number GOCE 037017, project OSIRIS. This work received also financial support from the Danish Council for Strategic Research, project REMTEC, contract 2104-07-0009. Support for this work was furthermore provided through a PhD grant of the University of Copenhagen for Charlotte N. Legind.

Model Availability Most models mentioned in this chapter are freely available via the authors. Please send an email if you are interested.

References

Alexander M (2000) Aging, bioavailability, and overestimation of risk from environmental pollutants. Environ Sci Technol 34:4259–4265

Baduru KK, Trapp S, Burken JG (2008) Direct measurement of VOC diffusivities in tree tissues: impacts on tree-based phytoremediation. Environ Sci Technol 42:1268–1275

Bakker M, Vorenhout M, Sijm DTHM, Kollöffel C (1999) Dry deposition of atmospheric polycyclic aromatic hydrocarbons in three plantago species. Environ Toxicol Chem 18, 2289–2294

Balouet J-C, Oudijk G, Smith KT, Petrisor I, Grudd H, Stocklassa B (2007) Applied dendroecology and environmental forensics. Characterizing and age dating environmental releases: fundamentals and case studies. Environ Forensics 8:1–17

Barret M (1995) Metabolism of herbicides by cytochrome P 450 in corn. Drug Metabol Drug Interactions 12:299–315

Boxall ABA, Johnson P, Smith EJ, Sinclair CJ, Stutt E, Levy LS (2006) Uptake of veterinary medicines from soils into plants. J Agric Food Chem 54(6):2288–2297

Börner H (1995) Unkrautbekämpfung. Gustav Fischer, Jena, D

Brand E, Otte PF, Lijzen JPA (2007) CSOIL 2000: an exposure model for human risk assessment of soil contamination. RIVM Report 711701054/2007, National Institute of Public Health and the Environment. http://rivm.openrepository.com/rivm/bitstream/10029/13385/1/711701054.pdf. Accessed 13 August 2010

Briggs GG, Bromilow RH, Evans AA (1982) Relationship between lipophilicity and root uptake and translocation of non-ionised chemicals by barley. Pestic Sci 13:495–504

Briggs GG, Bromilow RH, Evans AA, Williams M (1983) Relationships between lipophilicity and the distribution of non-ionised chemicals in barley shoots following uptake by the roots. Pestic Sci 14:492–500

Briggs GG, Rigitano RLO, Bromilow RH (1987) Physicochemical factors affecting uptake by roots and translocation to shoots of weak acids in barley. Pestic Sci 19:101–112

Buchholz A, Baur P, Schönherr J (1998) Differences among plane species in cuticular permeabilities and solute mobilities are not caused by differential size selectivities. Planta 206:322–328

Burken JG, Schnoor JL (1998) Predictive relationships for uptake of organic contaminants by hybrid poplar trees. Environ Sci Technol 32:3379–3385

Chamberlain K, Patel S, Bromilow RH (1998) Uptake by roots and translocation to shoots of two morpholine fungicides in barley. Pestic Sci 54:1–7

Chard BK, Doucette WJ, Chard JK, Bugbee B, Gorder K (2006) Trichloroethylene uptake by apple and peach trees: greenhouse study. Environ Sci Technol 40:4788–4793

Chiou CT, Sheng GY, Manes M (2001) A partition-limited model for the plant uptake of organic contaminants from soil and water. Environ Sci Technol 35(7):1437–1444

Collins CD, Bell JNB, Crews C (2000) Benzene accumulation in horticultural crops. Chemosphere 40:109–114

Cousins I, Mackay D (2001) Gas-particle partitioning of organic compounds and its interpretation using relative solubilities. Environ Sci Technol 35:643–647

Czub G, McLachlan MS (2004) A food chain model to predict the levels of lipophilic organic contaminants in humans. Environ Toxicol Chem 23:2356–2366

DEFRA (Department for Environment Food and Rural Affairs) (2002) The Contaminated land exposure assessment model (CLEA): technical basis and algorithms. R & D Publications CLR 10, Environment Agency. http://publications.environment-agency.gov.uk/pdf/SCLR10-e-p.pdf?lang=_e. Accessed 13 August 2010

Delschen T, Hembrock-Heger A, Necker U (1996) Systematische Untersuchungen zum Verhalten von PAK und PCB im System Boden/Pflanze auf der Lysimeter-Anlage Waldfeucht (1989–1994). In: Landesumweltamt NRW (eds) Materialien zur Ermittlung und Sanierung von Altlasten, LUA Essen, Germany, pp 1–214

Delschen T, Hembrock-Heger A, Leisner-Saaber J, Sopczak D (1999) Bedeutung verschiedener Belastungsursachen für den PAK-Gehalt von Kulturpflanzen. Umweltwissenschaften und Schadstoff-Forschung, UWSF – Z Umweltchem Ökotox 11:79–87

Dettenmaier EM, Doucette WJ, Bugbee B (2009) Chemical hydrophobicity and uptake by plant roots. Environ Sci Technol 43(2):324–329

Doucette WJ, Chard TJK, Moore BJ, Staudt WJ, Headley JV (2005) Uptake of sulfolane and diisopropanolamine (DIPA) by cattails (Typha latifolia). Microchem J 81:41–49

Dreicer M, Hakonson TE, White GC, Whicker FW (1984) Rainsplash as a mechanism for soil contamination of plant surfaces. Health Phys 46:177–187

Ebbs S, Bushey J, Poston S, Kosma D, Samiotakis M, Dzombak D (2003) Transport and metabolism of free cyanide and iron cyanide complexes by willow. Plant Cell Environ 26:1467–1478

EC (European Commission) (2003) Technical guidance document on risk assessment in support of commission directive 93/67/EEC on risk assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market, European Communities, Italy. http://ecb.jrc.it/documents/. Accessed 13 August 2010

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals) (2001) Exposure factors sourcebook for european populations (with focus on UK data). Technical Report No. 79, Brussels, Belgium

EFSA (European Food Safety Authority) (2006) Pesticides are EU citizens' top food-related health concern. Available online at http://www.euractiv.com. Accessed 17 March 2006

Fahl GM, Kreft L, Altenburger R, Faust M, Boedeker W, Grimme LH (1995) pH-Dependent sorption, bioconcentration and algal toxicity of sulfonylurea herbicides. Aquat Toxicol 31(2):175–187

Franco A, Trapp S (2008) Estimation of the soil-water partition coefficient normalized to organic carbon for ionizable organic chemicals. Environ Toxicol Chem 27(10):1995–2004

Franke W (1987) Nutzpflanzenkunde. Thieme Verlag, Stuttgart, Germany

Fryer ME, Collins CD (2003) Model intercomparison for the uptake of organic chemicals by plants. Environ Sci Technol 37:1617–1624

Fujisawa T, Ichise K, Fukushima M, Katagi T, Takimoto Y (2002a) Mathematical model for the uptake of non-ionized pesticides by edible root of root crops. J Pesticide Sci 27:242–248

Fujisawa T, Ichise K, Fukushima M, Katagi T, Takimoto Y (2002b) Improved uptake models of non-ionized pesticides to foliage and seed of crops. J Agric Food Chem 50: 532–537

Gent MPN, White JC, Parrish ZD, Isleyen M, Eitzer BD, Mattina MI (2007) Uptake and translocation of p,p′-dichlorodiphenyldichloroethylene supplied in hydroponics solution to Cucurbita. Environ Toxicol Chem 26(12):2467–2475

Gopalakrishnan G, Negri MC, Minsker BS, Werth CJ (2007) Monitoring subsurface contamination using tree branches. Ground Wat Monit Rem 27:65–74

Hsu FC, Marxmiller RL, Yang AYS (1990) Study on root uptake and xylem translocation of cinmethylin and related compounds in detopped soybean roots using a pressure chamber technique. Plant Physiol 93:1573–1578

Hülster A, Marschner H (1994) PCDD/PCDF-Transfer in Zucchini und Tomaten. Veröff PAÖ 8:579–589

Hülster A, Marschner H (1995) Identifizierung PCDD/PCDF-mobilisierender Verbindungen in Wurzelexudaten von Zucchini. Veröff PAÖ 12:359–369

Hülster A, Müller JF, Marschner H (1994) Soil-plant transfer of polychlorinated dibenzo-p-dioxins and dibenzofurans to vegetables of the cucumber family (Cucurbitaceae). Environ Sci Technol 28:1110–1115

Hung H, Mackay D (1997) A novel and simple model of the uptake of organic chemicals by vegetation from air and soil. Chemosphere 35:959–977

Inoue J, Chamberlain K, Bromilow RH (1998) Physicochemical factors affecting the uptake by roots and translocation to shoots of amine bases in barley. Pestic Sci 54:8–21

Kaupp H (1996) Atmosphärische Eintragswege und Verhalten von polychlorierten Dibenzo-p-dioxinen und -furanen sowie polyzyklischen Aromaten in einem Maisbestand. PhD thesis, faculty of biology, chemistry and earth science, University of Osnabrück

Kazerouni N, Sinha R, Hsu CH, Greenberg A, Rothman N (2001) Analysis of 200 food items for benzo[a]pyrene and estimation of its intake in an epidemiologic study. Food Chem Toxicol 39:423–436

Kleier DA (1988) Phloem mobility of xenobiotics. Plant Physiol 86:803–810

Komossa D, Langebartels C, Sandermann H Jr (1995) Metabolic processes for organic chemicals in plants. In: Trapp S, Mc Farlane C (eds) Plant contamination – modeling and simulation of organic chemical processes. Lewis, Boca Raton, Florida, pp 69–103

Kömp P, McLachlan M (1997) Interspecies variability of the plant/air partitioning of polychlorinated biphenyls. Environ Sci Technol 31:2944–2948

Kühne R, Ebert RU, Schüürmann G (2005) Prediction of the temperature dependency of Henry's Law constant from chemical structure. Environ Sci Technol 39:6705–6711

Larsen M, Trapp S (2006) Uptake of iron cyanide complexes into willow trees. Environ Sci Technol 40:1956–1961

Larsen M, Trapp S, Pirandello A (2004) Removal of cyanide by woody plants. Chemosphere 54:325–333

Larsen M, Ucisik A, Trapp S (2005) Uptake, metabolism, accumulation and toxicity of cyanide in willow trees. Environ Sci Technol 39:2135–2142

Larsen M, Burken J, Machackova J, Karlson UG, Trapp S (2008) Using tree core samples to monitor natural attenuation and plume distribution after a PCE spill. Environ Sci Technol 42:1711–1717

Legind CN, Trapp S (2009) Modeling the exposure of children and adults via diet to chemicals in the environment with crop-specific models. Environ Pollut 157:778–785

Lewis GN (1907) A new system of thermodynamic chemistry. Proc Am Acad Arts Sci 43: 259–297

Li JG, Gerzabek MH, Mück K (1994) An experimental study on mass loading of soil particles on plant surfaces. Bodenkultur 45:15–24

Lunney AI, Zeeb BA, Reimer K (2004) Uptake of weathered DDT in vascular plants: potential for phytoremediation. Environ Sci Technol 38:6147–6154

Mackay D, Fraser A (2000) Bioaccumulation of persistent organic chemicals: mechanisms and models. Environ Pollut 110:375–391

McKone TE, Maddalena RL (2007) Plant uptake of organic pollutants from soil: bioconcentration estimates based on models and experiments. Environ Toxicol Chem 26:2494–2504

Mikes O, Cupr P, Trapp S, Klanova J (2009) Uptake of polychlorinated biphenyls and organochlorine pesticides from soil and air into radishes (*Raphanus sativus*). Environ Pollut 157(2): 488–496

Monteith JL (1995) Accommodation between transpiring vegetation and the convective boundary-layer. J Hydrol 166(3–4):251–263

Müller JF, Hülster A, Päpke O, Ball M, Marschner H (1993) Transfer pathways of PCDD/F into fruits. Chemosphere 27:195–201

Müller JF, Hülster A, Päpke O, Ball M, Marschner H (1994) Transfer of PCDD/F from contaminated sites into carrots, lettuce and peas. Chemosphere 29:2175–2181

Mueller KE, Mueller-Spitz SR, Henry HF, Vonderheide AP, Soman RS, Kinkle BK, Shann JR (2006) Fate of pentabrominated diphenyl ethers in soil: abiotic sorption, plant uptake, and the impact of interspecific plant interactions. Environ Sci Technol 40:6662–6667

Paretzke HG, Garland JA (1990) Assessment of the radiological significance of surface contamination in entrained radioactivity. Final Report, EC-contract No 90-ET-015

Paterson S, Mackay D, McFarlane C (1994) A model of organic-chemical uptake by plants from soil and the atmosphere. Environ Sci Technol 28:2259–2266

Penman HL (1948) Natural evaporation from open water, bare soil and grass. Proc R Soc Lon A 193(1032):120–145

Pflugmacher S, Schröder P (1995) Glutathione S-transferases in trees: inducibility by various organic xenobiotics. Z Pflanzenernähr Bodenk 158:71–73

Reichenberg F, Mayer P (2006) Two complementary sides of bioavailability: accessibility and chemical activity of organic chemicals in sediments and soils. Environ Toxicol Chem 25:1239–1245

Riederer M (1995) Partitioning and transport of organic chemicals between the atmospheric environment and leaves. In: Trapp S, Mc Farlane JC (eds) Plant contamination. Modeling and simulation of organic chemical processes. Lewis, Boca Raton, FL, pp 153–190

Rikken MGJ, Lijzen JPA, Cornelese AA (2001) Evaluation of model concepts on human exposure. Proposals for updating the most relevant exposure routes of CSOIL. RIVM report 711 701 022, Bilthoven, NL

Russell RS, Shorrocks VM (1959) The relationship between transpiration and the absorption of inorganic ions by intact plants. J Exp Bot 10(29):301–316

Samiotakis M, Ebbs SD (2003) Possible evidence for transport of an iron cyanide complex by plants. Environ Pollut 127:169–173

Samsøe-Petersen L, Larsen EH, Larsen PB, Bruun P (2002) Uptake of trace elements and PAHs by fruit and vegetables from contaminated soils. Environ Sci Technol 36:3057–3063

Sandermann H (1994) Higher plant metabolism of xenobiotics: the 'green liver' concept. Pharmacogenetics 4:225–241

Sandermann H (2004) Bound and unextractable pesticidal plant residues: chemical characterization and consumer exposure. Pest Manag Sci 60:613–623

Satchivi NM, Stoller EW, Wax LM, Briskin DP (2000a) A nonlinear dynamic simulation model for xenobiotic transport and whole plant allocation following foliar application I. Conceptual foundation for model development. Pestic Biochem Phys 68:67–84

Satchivi NM, Stoller EW, Wax LM, Briskin DP (2000b) A nonlinear dynamic simulation model for xenobiotic transport and whole plant allocation following foliar application II. Model validation. Pestic Biochem Phys 68:85–95

SCF (Scientific Committee on Food) (2002) Annex: background document to the opinion of the scientific committee on food on the risks to human health of polycyclic aromatic hydrocarbons in food. Polycyclic aromatic hydrocarbons – occurrence in foods, dietary exposure and health effects. http://ec.europa.eu/food/fs/sc/scf/out154_en.pdf. Accessed 13 August 2010

Schönherr J, Riederer M (1989) Foliar penetration and accumulation of organic-chemicals in plant cuticles. Rev Environ Contam Toxicol 108:1–70

Shone MGT, Wood AV (1974) Comparison of uptake and translocation of some organic herbicides and a systemic fungicide by barley 1. Absorption in relation to physicochemical properties. J Exp Bot 25(85):390–400

Sicbaldi F, Sacchi GA, Trevisan M, Del Re AAM (1997) Root uptake and xylem translocation of pesticides from different chemical classes. Pestic Sci 50:111–119

Sitte P, Ziegler H, Ehrendorfer F, Bresinsky A (1991) Lehrbuch der Botanik für Hochschulen, 33rd edn. Gustav Fischer, Stuttgart

Sorek A, Atzmon N, Dahan O, Gerstl, Z, Kushisin, L, Laor Y, Mingelgrin U, Nasser A, Ronen D, Tschachansky L, Weisbrod N, Graber ER (2008) "Phytoscreening": the use of trees for discovering subsurface contamination by VOCs. Environ Sci Technol 42:536–542

Søvik AK, Alfnes E, Breedveld GD, French HK, Pedersen TS, Aagaard P (2002) Transport and degradation of toluene and o-xylene in an unsaturated soil with dipping sedimentary layers. J Environ Qual 31:1809–1823

TA-Luft (Technische Anleitung zur Reinhaltung der Luft) (1986) In: Handbuch des Umweltschutzes, Vol. II-2, appendix 3.1. ecomed, Landsberg aL, D

Ter Laak TL, Barendregt A, Hermens JLM (2006) Freely dissolved pore water concentrations and sorption coefficients of PAHs in spiked, aged, and field-contaminated soils. Environ Sci Technol 40:2184–2190

Thompson PL, Ramer LA, Guffey AP, Schnoor JL (1998) Decreased transpiration in poplar trees exposed to 2,4,6-trinitrotoluene. Environ Toxicol Chem 17:902–906

Trapp S (1995) Model for uptake of xenobiotics into plants. In: Trapp S, Mc Farlane JC (eds) Plant contamination. Modeling and simulation of organic chemical processes. Lewis, Boca Raton, FL, pp 107–152

Trapp S (2000) Modeling uptake into roots and subsequent translocation of neutral and ionisable organic compounds. Pest Manage Sci 56:767–778

Trapp S (2002) Dynamic root uptake model for neutral lipophilic organics. Environ Toxicol Chem 21:203–206

Trapp S (2004) Plant uptake and transport models for neutral and ionic chemicals. Environ Sci Pollut Res 11:33–39

Trapp S (2007) Fruit Tree model for uptake of organic compounds from soil and air. SAR QSAR Environ Res 18:367–387

Trapp S, Christiansen H (2003) Phytoremediation of cyanide-polluted soils. In: McCutcheon SC, Schnoor JL (eds) Phytoremediation: transformation and control of contaminants. Wiley, Hoboken, NJ, pp 829–862

Trapp S, Karlson U (2001) Aspects of phytoremediation of organic compounds. J Soils Sed 1: 37–43

Trapp S, Matthies M (1995) Generic one-compartment model for uptake of organic chemicals by foliar vegetation. Environ Sci Technol 29:2333–2338; erratum 30, 360

Trapp S, Pussemier L (1991) Model calculations and measurements of uptake and translocation of carbamates by bean plants. Chemosphere 22:327–339

Trapp S, Schwartz S (2000) Proposals to overcome limitations in the EU chemical risk assessment scheme. Chemosphere 41:965–971

Trapp S, Matthies M, Scheunert I, Topp EM (1990) Modeling the bioconcentration of organic chemicals in plants. Environ Sci Technol 24:1246–1252

Trapp S, Matthies M, Mc Farlane C (1994) Model for uptake of xenobiotics into plants: validation with bromacil experiments. Environ Toxicol Chem 13:413–422

Trapp S, Yu X, Mosbæk H (2003) Persistence of methyl tertiary butyl ether (MTBE) against metabolism by Danish vegetation. Environ Sci Pollut Res 10:357–360

Trapp S, Cammarano A, Capri E, Reichenberg F, Mayer P (2007a) Diffusion of PAH in potato and carrot slices and application for a potato model. Environ Sci Technol 41:3103–3108

Trapp S, Ucisik AS, DelChicca Romano P, Larsen M (2007b) The role of plants and bacteria in phytoremediation – kinetic aspects. In: Heipieper HJ (ed) Bioremediation of soils contaminated with aromatic compounds. NATO science series, IV. Earth and environmental sciences – vol 76. Springer, Dordrecht, NL, pp 41–49

Trapp S, Feificova D, Rasmussen NF, Bauer-Gottwein P (2008) Plant uptake of NaCl in relation to enzyme kinetics and toxic effects. Environ Exp Bot 64: 1–7

Trapp S, Larsen M, Legind CN, Burken J, Machackova J, Karlson UG (2009) A guide to vegetation sampling for screening of subsurface pollution. Freely available at http://homepage.env.dtu.dk/stt/. Accessed 13 August 2010

Travis CC, Arms AD (1988) Bioconcentration of organics in beef, milk, and vegetation. Environ Sci Technol 22:271–274

Travis CC, Hattemer-Frey HA (1991) Human exposure to dioxin. Sci Total Environ 104:97–127

Ucisik AS, Trapp S (2006) Uptake, removal, accumulation, and phytotoxicity of phenol in willow trees (*Salix viminalis*). Environ Toxicol Chem 25:2455–2460

Ucisik AS, Trapp S (2008) Uptake, accumulation, phytotoxicity and removal of 4-chlorophenol in willow trees. Arch Environ Contam Toxicol 54:619–627

Ucisik AS, Trapp S, Kusk KO (2007) Uptake, accumulation, phytotoxicity and removal of 2,4-dichlorophenol in willow trees. Environ Toxicol Chem 26:1165–1171

Vroblesky DA, Nietch CT, Morris JT (1999) Chlorinated ethenes from groundwater in tree trunks. Environ Sci Technol 33:510–515

Whitfield Åslund ML, Zeeb BA, Rutter A, Reimer KJ (2007) In situ phytoextraction of polychlorinated biphenyl – (PCB) contaminated soil. Sci Total Environ 374:1–12

White JC (2002) Differential bioavailability of field-weathered p,p′-DDE to plants of the Cucurbita and Cucumis genera. Chemosphere 49:143–152

Yu X, Trapp S, Puhua Z, Chang W, Xishi Z (2004) Metabolism of cyanide by Chinese vegetation. Chemosphere 56:121–126

Yu X, Trapp S, Zhou P, Hu H (2005a) The effect of temperature on the rate of cyanide metabolism of two woody plants. Chemosphere 59:1099–1104

Yu XZ, Trapp S, Zhou P (2005b) Phytotoxicity of cyanide to weeping willow trees. Environ Sci Pollut Res 12:109–113

Yu XZ, Trapp S, Zhou PH, Chen L (2007) Effect of temperature on the uptake and metabolism of cyanide by weeping willows. Int J Phytorem 9:243–255

Zhu Y, Zhang S, Zhu Y-G, Christie P, Shan X (2007) Improved approaches for modeling the sorption of phenanthrene by a range of plant species. Environ Sci Technol 41:7818–7823

Zohair A, Salim A-B, Soyibo AA, Beck AJ (2006) Residues of polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs) and organochlorine pesticides in organically-farmed vegetables. Chemosphere 63:541–553

Chapter 10
Vapor Intrusion

Todd A. McAlary, Jeroen Provoost, and Helen E. Dawson

Abstract Vapor intrusion is a pathway of potential exposure to *volatile* and *semi-volatile contaminants* (collectively referred to here as VOCs or vapors) that migrate from the subsurface to the air inside occupied buildings. Soil vapor intrusion to indoor air can occur regardless of whether a building has a basement, slab-on-grade or crawlspace design. As a basis for a mathematical model a c*onceptual model* is needed, which describes the movement of contaminants from the source to the building, vapor migration barriers and receptors. It also provides a framework for interpreting the processes influencing the fate and transport of contaminants as they move from a source to a receptor. The approach for assessing vapor intrusion will vary from site-to-site, but there are certain elements that are appropriate in most cases. In this chapter the relevant processes have been described, like phase partitioning, biodegradation, advection and dilution within the building due to ventilation. Moreover, the presence of NAPLs, available vapor intrusion models, sampling and analysis procedures and subsurface vapor mitigation have been discussed.

Contents

T.A. McAlary (✉)
Geosyntec Consultants, Inc., Guelph, Canada
e-mail: tmcalary@geosyntec.com

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_10,

10.1 Introduction

Vapor intrusion is a pathway of potential exposure to *volatile* and *semi-volatile contaminants* (collectively referred to here as VOCs or vapors) that migrate from the subsurface to the air inside occupied buildings. This particular pathway has gained increasing attention in the past decade because of several highly publicized cases where volatile contaminants were detected in indoor air samples at concentrations higher than human health risk-based quality standards. To date, the practical experience with assessing vapor intrusion has demonstrated that this pathway can be very challenging to assess for several reasons. The risk-based quality standards for indoor air are very low, so extra care is required to avoid positive bias from contaminated sampling equipment. Several common contaminants are typically found in indoor air in residences at concentrations at or above risk-based quality standards due to emissions from consumer products, building materials and even outdoor air sources, so the relative contribution from the subsurface is often difficult to resolve. Indoor air concentrations vary with the weather and building ventilation rate, and concentrations of VOCs in soil gas and groundwater samples show spatial and temporal variability. All of these issues contribute to a relatively high level of uncertainty in the assessment of vapor intrusion.

This chapter describes the factors affecting the vapor intrusion pathway, provides information that can help in assessment of the pathway, and includes references to much more detailed literature on the topic.

10.2 Conceptual Models

A *conceptual model* provides a framework for interpreting the processes influencing the fate and transport of contaminants as they move from a source to a receptor. A conceptual model is not the same as a mathematical model, but should be the

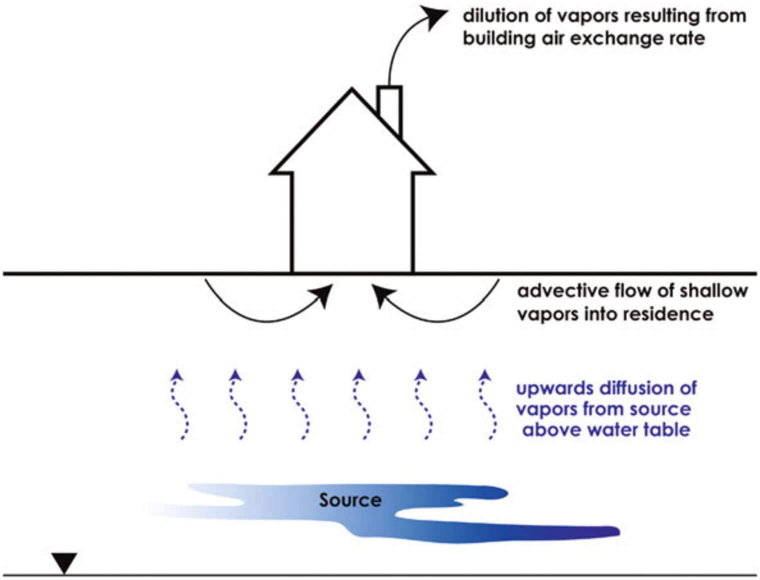


Fig. 10.1 Simplified conceptual model of vapor intrusion

basis for one. A conceptual model combines available site-specific information with theoretical considerations and experience from similar sites to form an expectation of site conditions. As site-specific data are collected, they are compared with the conceptual model and revisions are made as needed. Conceptual models are often described using a source-pathway-receptor framework, which has been adopted herein. Thus, the conceptual model evolves over time as information is collected, becoming more comprehensive, and sometimes more complex in the process.

The most common conceptual model for subsurface vapor intrusion considers a source of contaminants at some depth below a building, upward diffusion of vapors at steady-state, convection into the building from the close proximity of the floor slab, and dilution within the building due to ventilation. This simplified conceptual model is shown schematically in Fig. 10.1. However, several other mechanisms may be important, depending on site-specific conditions, as discussed further below.

10.2.1 *Vapor Source*

The location of the vapor source relative to the water table or capillary fringe can be important when assessing vapor intrusion at a particular site. In general, a source can exist above the saturated zone, just above the water table (e.g., non-aqueous phase liquid (NAPL) floating on the water table or disseminated within the capillary fringe) or in the saturated zone (e.g., a soluble groundwater plume).

The approximate mass of the source should be considered, to assess whether it is large enough to be a persistent source for a typical exposure scenario (for about 30 years), or not. A very small release can result in highly localized soil vapor concentrations that may result in (temporary) high concentrations in the indoor air, but may not be sufficient to sustain an unacceptable flux to indoor air over a 30-year exposure period. Estimating the mass of a source is challenging, and the uncertainty in the estimate should be considered as part of the conceptual model.

The source constituents (i.e. the contaminants in the soil and/or groundwater) should be evaluated and identified to the extent practical. A solvent release may be predominantly a single contaminant, although intrinsic biodegradation may generate daughter products that may be more or less mobile and/or toxic. By contrast, a hydrocarbon release typically contains a mixture of hundreds of contaminants, of which many may be non-toxic and readily degradable under aerobic conditions, and only very few may be significant contributors to health risk. Waste disposal or recycling facilities may have even more complex mixtures. For NAPL mixtures, volatilization of individual contaminants is dependent on the composition of the mixture and the fraction of each contaminant may change significantly as weathering proceeds. Typically, lighter contaminants are more volatile and will decrease in concentration with time, leaving a mixture with increasing relative proportions of heavier, less volatile contaminants. The volatility, toxicity, mobility, degradability and initial mass fraction may all need to be considered for multiple contaminants to identify primary contaminants of concern and develop a practical list of target contaminant to be included in sampling and analysis.

Some sources produce vapors only by volatilization from aqueous or non-aqueous releases, but others actively produce gases such as methane, carbon dioxide or other volatile metabolites as products of biodegradation. Gas produced as a byproduct of microbiological activity can generate pressure gradients that enhance subsurface vapor migration by advection, in addition to diffusion. Methane is not usually a threat to indoor air quality, but in some circumstances, the indoor air concentration may exceed the lower explosive limit (5% by volume). Since methane cannot be present in soil gas at concentrations above 100%, this circumstance can only occur where the building air exchange rate is low, and is often associated with falling barometric pressure, which can increase the flow of soil gas into a building.

10.2.2 Pathway

The movement of contaminants from the source to the building is described in the pathway component of the conceptual model. In some settings, the pathway might include, for example, soil gas transport from a source above the water table to an overlying building. In other settings, the pathway might include a combination of groundwater transport from a primary source to an adjacent property, with subsequent off-gassing and vapor transport through the water-unsaturated zone to an overlying building. Alternately, the pathway may include lateral diffusion through

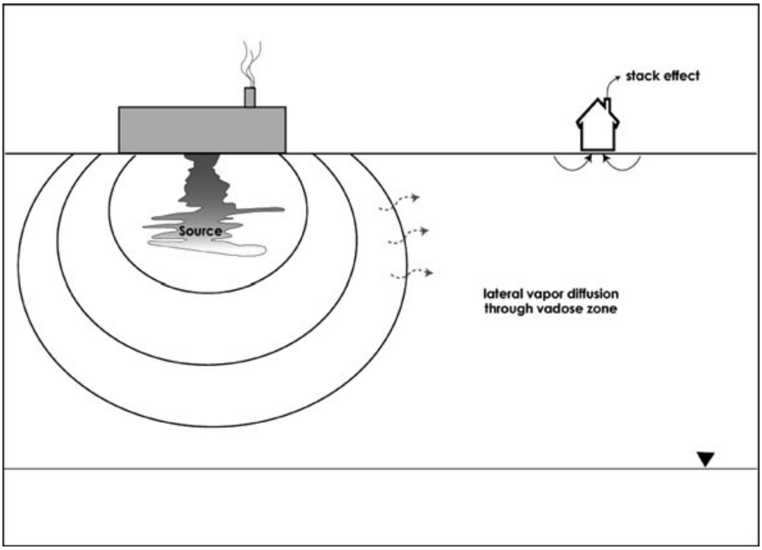


Fig. 10.2 Lateral diffusion of vapors through the water-unsaturated zone (schematic)

the water-unsaturated zone without contact or interaction with groundwater, as shown in Fig. 10.2.

There may also be preferential soil gas flow through granular fill under a building (Fig. 10.3), especially in areas where the gas permeability of the surrounding soil is very low. Floor drains, for example, are designed to allow water to drain away, but are usually not designed or constructed to eliminate soil gas entry. The granular materials surrounding a sewer pipe may or may not be very well compacted after placement, so settlement over time may form air voids beneath the slab that are very highly permeable. Foundation walls are usually constructed first, then floor slabs are poured, often leaving a space between the floor slab and walls (i.e., perimeter crack) for expansion and contraction. This perimeter crack is often obscured by wall-coverings, and may not be accessible for inspection or direct testing. The same may be true for other utility penetrations and homeowner modifications, which may also create a pneumatic connection to granular fill.

Sumps or wet basements can also allow groundwater containing VOCs to enter the building, or contact the building envelope directly (Fig. 10.4). This scenario can be especially problematic if the source of vapors is in the form of low-density non-aqueous phase liquids (LNAPLs) floating on the water table.

Conceptualization of the pathway should also include potential vapor migration barriers. A low-permeability layer in the water-unsaturated zone with high moisture content or perched water may impede or prevent upward migration of vapors from deeper sources. If the recharge rate is sufficient to cause perched water layers within the water-unsaturated zone, there may be an effective vapor barrier. The seasonality of infiltration would need to be considered in this scenario. Areas that receive regular

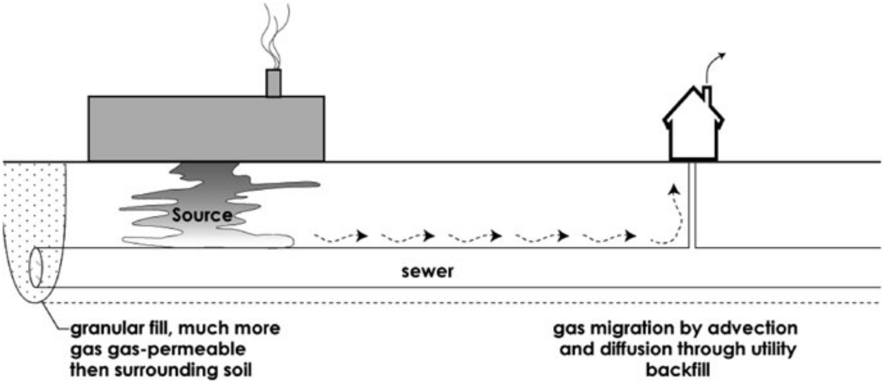


Fig. 10.3 Vapor transport through preferential pathways (schematic)

rainfall will be more likely to sustain layers of high moisture content in the water-unsaturated zone than areas with wet and dry seasons. Also, there may be a "rain-shadow" beneath a building where moisture contents are lower. This may depend on soil type and the size of the building.

In humid climatic regions and areas with artificial recharge (e.g., irrigation, storm-water retention ponds), a layer of clean water may accumulate above a plume of VOCs in groundwater, the thickness of which would typically grow with increasing travel distance and time from the point of release. This condition has been referred to as a "fresh-water lens" (Fitzpatrick and Fitzgerald 1996, 2002; McAlary et al. 2004) or a "diving plume" (API 2006), and can act as an effective vapor barrier, inhibiting off-gassing of VOCs from the water table sufficiently to protect overlying

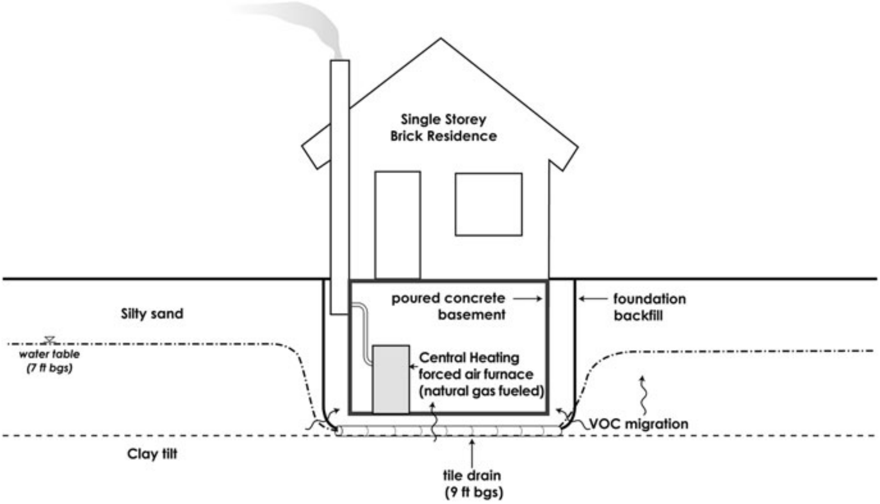


Fig. 10.4 Vapor intrusion in a building with a wet basement (schematic)

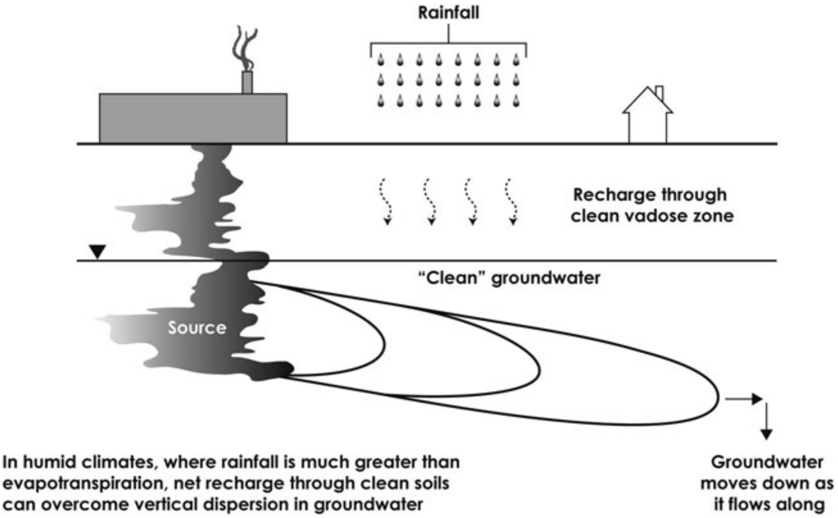


Fig. 10.5 Fresh water lens as a barrier to volatilization from groundwater (schematic)

buildings from subsurface vapor intrusion, as shown in Fig. 10.5. This condition may not develop if water table fluctuations are large, which would result in vertical mixing of VOCs across the upper water-saturated zone.

In some cases, geologic layers (soil or rock) can impose sufficient resistance to vapor transport to make the vapor intrusion pathway insignificant, providing the geologic features are laterally extensive over distances that are large compared to the size of the building or contaminant distribution. Characterization of the soil and extent to which specific soil layers may act as a barrier to vapor transport may require investigative methods such as pneumatic or geophysical testing, which are different from the sampling and analysis techniques commonly used to assess the vapor intrusion pathway. Nevertheless, the physical data can in some cases provide a definitive line of evidence, depending on the site conditions.

10.2.3 Receptor

For vapor intrusion studies, the receptor is usually the occupant(s) of a building. Receptors are generally either residents or workers, with exposure frequency and duration, and possibly inhalation rate being the primary differences in the exposure scenarios. Resident or worker exposures may be compared to risk-based targets or to indoor air quality standards specified by the regulatory agency with jurisdiction in the locale in which the assessment is conducted. The applicable indoor air quality standard should be defined as early as possible in the vapor intrusion pathway assessment process.

In some circumstances, the building itself could be considered a "receptor". Several aspects of the building design, build quality and ventilation characteristics can have significant influences on vapor intrusion. Where buildings are heated, convection cells develop with hot air rising and leaking through roofs and upper-floor windows. This phenomenon is referred to as the “stack effect”, and is commonly assumed to cause de-pressurization in the bottom floor of a building by a few Pascals. The “stack-effect” is a function of the height of a building and the temperature difference between inside and outside, and can result in under-pressurization of hundreds of Pascals in very tall buildings in cold climates. The escaping air will be replaced to some extent by soil gas entry at lower levels/basements. Even where buildings are not mechanically heated, solar radiation on rooftops can heat air in the attic significantly and cause the same effect. Wind-load on the side of a building can also cause a pressure differential across the building, and wind flow over the building can create a vacuum within the building via Bernoulli’s principle. Bathroom exhaust fans, central vacuum cleaners, clothes dryers and kitchen exhaust fans all remove air from a building. Elevator shafts can also cause localized pressure gradients. Window-mounted air conditioners blow air into a building. In commercial/industrial buildings, heating, ventilation and air conditioning (HVAC) units are usually mounted on the roof, and blow air into the building, heating or cooling it as appropriate for the season or climate. These units may also be designed to exhaust a portion of the indoor air to provide a certain amount of outdoor air into the building as fresh air or “make-up air”. Operation of HVAC units can generate sufficient pressure or vacuum to significantly influence vapor intrusion, and may be worth considering as part of the conceptual model. Verification of these pressure gradients typically requires a digital micro-manometer, which are in common use in the HVAC industry. In many cases, valuable information like air exchange rate, building pressure, seasonal changes, et cetera, can be obtained from the HVAC engineer, with minimal effort.

10.2.4 Vapor Intrusion Assessment Approach

The approach for assessing vapor intrusion will vary from site-to-site, but there are certain elements that are appropriate in most cases (see Fig. 10.6). It is usually advisable to begin a vapor intrusion assessment with a thorough planning step. This step typically includes gathering readily available existing information, formulating an initial conceptual model, establishing a scope for an initial phase of investigation, and developing a logical plan for future directions in response to the probable range of outcomes. Having a clear and logical plan will facilitate communication with building owners, occupants, regulators and other stakeholders. The plan should include a rationale or logic for how the data will be interpreted, including the basis for the indoor air quality target, any threshold above which an interim action will be required, whether confirmatory sampling will be required if concentrations are all below target levels, and similar considerations.

In some cases it may be appropriate to conduct a focused assessment of one particular aspect of the pathway, for example at sites where there is reason to believe that a particular process is effectively preventing subsurface vapor intrusion, and a

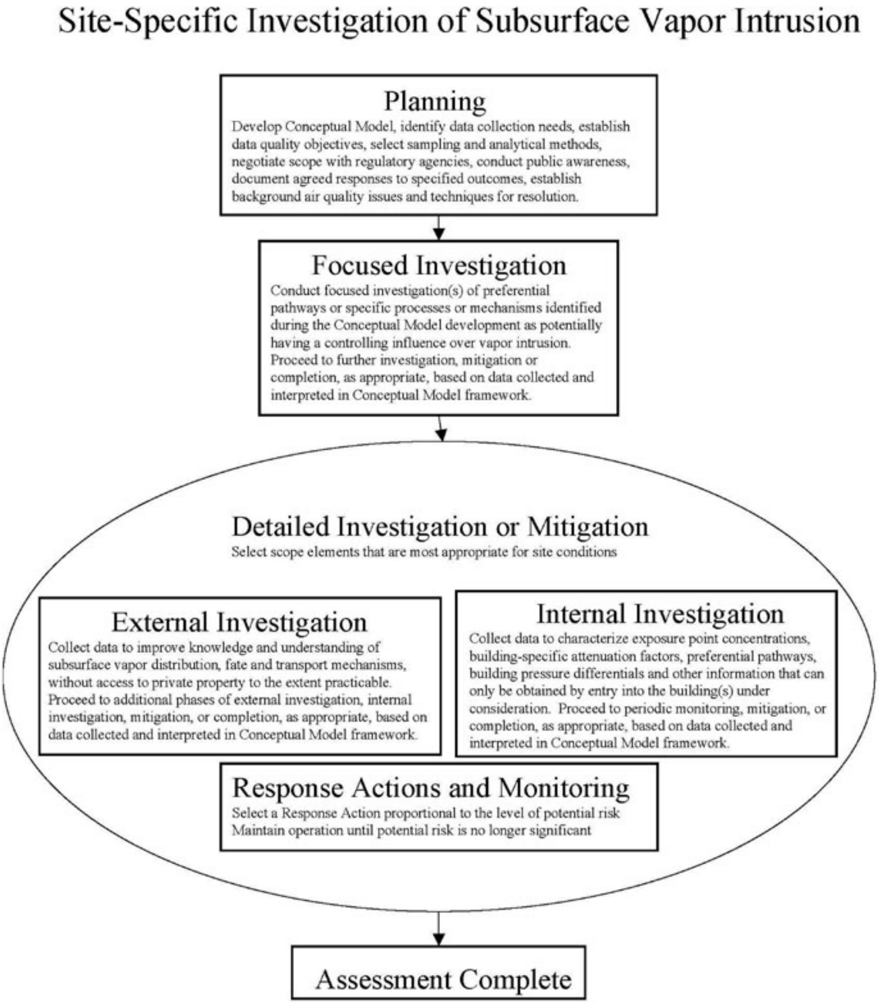


Fig. 10.6 Generic vapor intrusion assessment approach

limited scope of work can be conducted to demonstrate this with sufficient confidence. Even where a focused investigation may not be conclusive, it may still yield valuable information for scoping subsequent phases.

If further investigation is justified, an appropriate first step would include characterizing soil gas concentrations in proximity to the known or suspected sources of vapors. If such a survey identifies concentrations of potential concern, additional activity may be appropriate to delineate the extent of soil vapors. However, if source concentrations are too low to pose a concern for indoor air quality, the scope of any further phases or confirmatory monitoring may be curtailed appropriately, or the Risk Assessment even could be concluded. Source area investigations should generally be designed to provide basic information regarding the contaminants present, their relative concentrations, possible presence (and distribution, if possible) of NAPL, and temporal variability through periodic monitoring in selected locations. In addition to sampling for chemical analysis, it is advisable to collect other data for use in screening level modeling, e.g., permeability, moisture content, biodegradation assessment information. The site conceptual model should be updated as appropriate after these additional data are collected.

If site-specific assessments of the source and distribution of soil vapors indicate the potential for vapor intrusion, further characterization may be appropriate that requires a sampling team to enter a building (i.e., "internal investigations", which typically include indoor air and sub-slab soil gas sampling and analysis). While it may be tempting to conduct internal investigations as early as possible in the process, experience has shown that in most cases, indoor air samples contain detectable concentrations of VOCs from consumer products, building materials and even outdoor air that are within the range of risk-based target concentrations for several compounds (Dawson and McAlary 2008), in which case, it can be difficult or impossible to interpret the relative contributions from subsurface vapor intrusion. Subsurface samples (soil gas or groundwater) are less likely to be significantly affected by background contributions, and the relative proportions of the contaminants in the subsurface can be compared to the relative proportions of the same contaminants in indoor air to assess consistency (indicating vapor intrusion is the dominant contribution to indoor air VOCs) or inconsistency (indicating background contributions are likely significant). Thus, it is usually valuable to have a reasonable amount of data from assessing the source and subsurface pathway before conducting internal investigations.

If the distribution of subsurface vapors has been adequately mapped, there may be sufficient information with which to justify a "primary" investigative zone (most likely to have vapor intrusion at unacceptable levels), a "secondary" zone (unlikely to have unacceptable vapor intrusion, but included in the monitoring program as a precaution to account for spatial variability and subsurface heterogeneity), and a "tertiary" zone which would not be monitored unless nearby properties in the secondary zone are found to have unacceptable vapor intrusion. In some cases, it may be more practical to proceed with implementing exposure controls in lieu of further investigation. It is important to monitor the effectiveness of mitigation measures since soil air migration might change over time. More information about vapor intrusion mitigation can be found in EPA (1993b), Eklund et al. (2007), Folkes (2002) and ITRC (2007).

10.3 Fate and Transport Processes

Knowledge of the theory of vapor fate and transport mechanisms is essential for interpreting the data collected during a site-specific assessment of subsurface vapor intrusion. It is also invaluable for selecting data collection activities to focus on the most important locations, contaminants, or parameters in order to maximize the data value. Theoretical considerations should be the backbone of the conceptual model and any mathematical models. Data will help to identify "what" is happening, but theory is required to understand "why".

Mass transfer of VOCs from water to air has been extensively studied for surface water, but less so for groundwater. In either type of system, mass transfer is viewed as occurring over four layers: bulk liquid, liquid at a boundary layer between the two phases, air at the boundary layer, and bulk air. Mass transfer occurs by transport from the bulk liquid to liquid at a boundary layer, phase transfer to air at the boundary layer, and then transfer of the gaseous constituent from the air boundary layer to the bulk atmosphere. The overall mass transfer rate is limited by the slowest step in the process. In above ground systems, this typically is the transfer from the bulk liquid to the liquid at the boundary layer. In subsurface systems, however, there are additional considerations, such as vapor transport through water-unsaturated porous media (including the capillary fringe above a groundwater table), phase partitioning during transport, and biodegradation as described in more detail below.

10.3.1 Phase Partitioning

Phase partitioning calculations can be performed to estimate the concentration in any one of the phases (soil gas, pore water or solid phase) from the concentration in another phase (i.e., a soil gas sample), or the sum of all phases (i.e., a bulk soil sample), providing the total pore water content and fraction of organic carbon (Foc) of the soil are known. These soil properties and bulk soil VOC concentrations can be highly variable on small scales (i.e., the scale of typical soil sample volumes), so partitioning calculations generally reflect this variability (Davis et al. 2005; Tillman and Weaver 2007a). Therefore, it is generally best not to try to calculate soil gas concentrations from bulk soil concentrations, but rather to measure soil gas concentrations directly (Provoost et al. 2009). Many regulatory guidance documents do not consider screening the vapor intrusion pathway using soil VOC concentrations for this reason.

The maximum soil vapor concentration that can be achieved in immediate proximity to a NAPL can be calculated as the ratio of the vapor pressure to the total pressure, which is essentially atmospheric pressure. This can be thought of as a saturation limit in air. If the NAPL consists of a mixture of contaminants, Raoult's Law can be used to calculate the maximum soil vapor concentration, which requires measurement or estimation of the mole fractions of the contaminants in the mixture. Depending on the complexity of the mixture, variations from this ideal behavior may be significant.

Measured groundwater and soil gas concentrations do not always show the ratio predicted by Henry's Law (Provoost et al. 2008). This is partly because groundwater samples tend to be a mixture of water from different depths along the screened interval instead of being only from just below the water table, and it is also partly because of rate-limiting mass transfer across the capillary fringe and the fact that the subsurface is not a closed system. Therefore, it is usually advisable to collect deep soil gas data to provide quantitative evidence of the degree of off-gassing from the water table where groundwater concentrations appear to be high enough to cause a potential vapor intrusion concern. Contaminant transport from groundwater to the water-unsaturated zone can occur by diffusion through the capillary fringe (i.e., tension-saturated zone), but the diffusion coefficient through the water-saturated capillary fringe will be low. The capillary fringe thickness depends on the texture of the soil, ranging from a few centimeters in sands and gravels (which may impose negligible resistance to off-gassing) to many feet/meters for clays. Mass transfer to the water-unsaturated zone will increase if the water table fluctuates, because a falling water table leaves impacted water above the water table in tension that was previously below the water table and thereby facilitates off-gassing (McCarthy and Johnson 1991). This is shown in Fig. 10.7. Mass transfer to the water unsaturated zone will be reduced if a high rate of rainfall recharge creates a layer of clean water at the water table (called a fresh-water lens), providing that the infiltration does not occur through a NAPL-impacted vadose zone or one containing VOC vapors. This layer may act as an effective barrier to off-gassing from the water table as shown in Fig. 10.5. Many contaminated sites have groundwater monitoring wells with many years of water level monitoring data that can quickly be reviewed to assess the potential for each of these conditions.

Transport from deep soil gas to buildings typically is dominated by diffusion. Molecular diffusion results in movement of contaminants away from areas of higher concentrations toward areas of lower concentration. As soil is porous, with some of the pore space air-filled and some water-filled, the overall effective vapor-phase diffusion coefficient for a contaminant through soil can be estimated as a combination of diffusion through the soil vapor and the water in the pores. Free air diffusion coefficients are typically about 10,000 times higher than aqueous diffusion

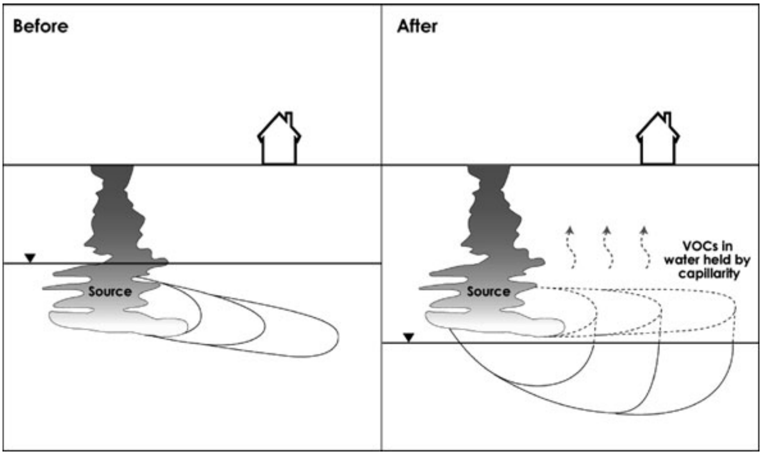


Fig. 10.7 Effect of falling water table on off-gassing from groundwater (schematic) vapor diffusion in soils

coefficients, so the diffusive flux in the aqueous phase will generally be negligible (except for compounds with very low Henry's Law Constants, where the mass fraction in the gas phase is very small). Thus, effective diffusion coefficients in soil tend to be high, although very dry soil can also result in low effective diffusion coefficients by direct mineral sorption of contaminants to the solid phase (McAlary 1989). Under nearly water-saturated conditions, such as within the capillary fringe, the effective diffusion coefficient may be quite low. As a consequence, transport through the capillary fringe is often the limiting factor in vapor transport from groundwater into indoor spaces. Contaminants with very low Henry's Law Constants (<0.001) may also have very low effective diffusion coefficients.

10.3.2 Biodegradation

Several vapor intrusion studies (DeVaull et al. 2002; Fischer et al. 1996; Hers et al. 2000; Lahvis et al. 1999; Ostendorf and Kampbell 1991; Roggemans 1998; Roggemans et al. 2001) have demonstrated that the aerobic biodegradation of hydrocarbon vapors can be significant in the vadose zone. These studies reached these conclusions through:

- field investigations examining soil gas concentration profiles of petroleum hydrocarbons and indicators of biological activity (oxygen and carbon dioxide);
- a comparison of modeled and measured vapor intrusion attenuation factors for petroleum and non-petroleum contaminants;
- modeling studies to characterize the potential impact of biodegradation on the indoor air concentrations.

There is uncertainty regarding whether and to what extent the presence of a building inhibits the supply of oxygen and therefore, whether degradation below a building is similar to degradation beside a building. Few studies to date have generated data specifically to assess the extent of degradation directly beneath buildings (Luo et al. 2009; McAlary et al. 2007).

DeVaull et al. (1997) listed conditional criteria for aerobic biodegradation of aromatic hydrocarbons in soil. In order for biodegradation to occur, sufficient hydrocarbon, oxygen, nutrients, moisture, and microbial populations must be present. Typically, sufficient oxygen is the limiting criteria since sufficient microbes, soil moisture, and nutrients are present at most sites. Biodegradation of contaminant vapors can potentially occur by anaerobic processes where oxygen has been consumed or by cometabolic processes where there are appropriate mixtures of primary metabolites and cometabolites, however, both of these processes tend to be much less significant than aerobic metabolism and have not been demonstrated for vadose zone vapor transport.

The rate and extent of aerobic degradation varies from site to site (Roggemans et al. 2001), such that it is difficult to predict the degree of degradation a-priori, or even model it without collecting some site-specific information regarding the vertical profile of O_2, CO_2 and VOCs. Methods have been developed to model vadose zone biodegradation beneath open surfaces (Lahvis and Baehr 1997 and 1996) and beneath buildings (Abreu and Johnson 2005, 2006; Johnson and Abreu 2003, Johnson et al. 1999). First order and Monod degradation kinetics have been included in these models. Some of the more refined methods model the coupled transport and reaction of hydrocarbons and oxygen because depletion of oxygen in the subsurface may result from consumption resulting from the biodegradation process.

10.3.3 Soil Gas Advection

Advection of soil gas into buildings occurs due to under-pressurization of the building relative to pressure in shallow soil gas. The building under-pressurization can be a result of the “stack effect” (warm air rising inside the building, exiting through roof vents, and creating suction in lower levels), barometric pumping, wind load on the side of the building, exhaust fans, clothes dryers, central vacuums, or elevators (acting as a piston). The soil gas flow rate into a building is a function of building pressure, permeability of soils immediately beneath the foundation, and characteristics of vapor entry points through the foundation.

The pressure differential between the building and the subsurface will often fluctuate between positive and negative, in which case air will flow into and out of the building, which can affect the soil gas in proximity to points of entry to the building (e.g., floor cracks, drains, et cetera). This can result in dilution of sub-slab concentrations, contribution of vapors from indoor sources to the sub-slab soil gas, and/or addition of oxygen to the sub-surface, which may have a significant effect on biodegradation of petroleum hydrocarbons. Pressure differences may also result in indoor air flowing to the soil air as McHugh et al. (2006) demonstrated. Monitoring the pressure differential with a digital micro-manometer and datalogger can help to understand whether and to what effect pressure fluctuations are important. In some cases, indoor air monitoring under positive and negative building pressure can help to elucidate the contribution of vapors from the subsurface (Berry-Spark et al. 2005).

Soil vapor intrusion to indoor air can occur regardless of whether a building has a basement, slab-on-grade or crawlspace design. It can also occur even when the building has a concrete floor that appears to be free of cracks. Vapor barriers are placed under some buildings to attempt to inhibit subsurface vapor intrusion, but the barriers are typically plastic sheets that can reduce advection of soil gas, but may not significantly impede diffusion. If not properly designed and installed, vapor barriers may contain inadvertent perforations, and air flow may be significant even through a small perforation. Intrinsically safe buildings may be limited to those constructed on stilts, such as may occur in lowland areas subject to regular flooding.

10.3.4 Mixing Inside the Building

Once subsurface vapors migrate to the building, air circulation within the building will result in a mixing of the contaminant through the interior. Air exchange rates within buildings depends on the design of the building heating and ventilation systems and can range from 0.1 to 10 per h. (ASHRAE 1985). In some cases, the air exchange rate is specified by building codes, or other standards (e.g., ANSI/ASHRAE 2004). Forced-air systems generally circulate air at a rate sufficient to maintain well-mixed conditions within a given air-zone (building or part of a building which is serviced by a single ventilation system). Buildings heated with baseboard heaters or radiators tend to have less circulation, limited to convection and currents imposed by wind-load on the building, leaking doors, and windows and other openings. In very large buildings, there are often a series of air-handling units, and different air-zones, which may be well-mixed internally, but may be relatively isolated from one another. Considering the range of possible configurations, there is a knowledge gap related to assigning a representative number of indoor air samples as a function of the building design and ventilation system design.

The characterization and modeling of contaminants in indoor air has been extensively studied in the fields of industrial hygiene and energy efficiency, but vapor intrusion guidance documents typically assume a very simple scenario where the entire indoor space is perfectly and instantaneously well-mixed and data from basement levels (or even crawl spaces beneath buildings) represents potential exposures in living areas. In reality, concentrations will be higher in rooms with limited ventilation (e.g., cellars) or immediately adjacent to points of entry (sumps, floor cracks, et cetera) to the point where concentrations may be detectable with field instruments. In large buildings, the ventilation system design and operation data may be available from mechanical engineers or building custodians, but in most single-family dwellings, the ventilation rate and mixing between air-spaces will be unknown.

10.4 Mathematical Modeling

Mathematical modeling can help develop an understanding of the significance of various fate and transport processes, compare expected performance of various remedy designs (if needed), and possibly act as an additional line of evidence or interpretive tool. In general, a mathematical model of vapor intrusion predicts the indoor air concentration that may result through vapor transport from a subsurface contaminant source into a building. Screening level models for a given set of site conditions can be performed quickly, so it is usually appropriate to conduct a few simple bounding calculations to assess “best-case” and “worst-case” conditions. If vapor intrusion is significant in both cases, it may be preferable to proceed toward remedy design than an expensive investigation. Conversely, if vapor intrusion is insignificant in both cases, it may be possible to select a focused scope of investigation, and use the model calculations as a supporting line of evidence. If appropriate input data are available, screening level models will generally provide results that are similar to measured conditions, but considering the spatial, temporal, sampling and analytical variations, uncertainty in model inputs, and limitations of simplistic model formulations, it should not be surprising if measurements and model outputs show discrepancies of approximately one order of magnitude (Hers et al. 2000; Johnson 2002; Johnson et al. 2002; Tillman and Weaver 2007b).

10.4.1 Mathematical Model Formulation

As illustrated in the conceptual model, the vapor intrusion pathway commonly involves a phase partitioning component (transfer of contaminants from a source to the soil gas phase), diffusive vapor transport through the water-unsaturated zone, convective/diffusive transport through a building's foundation, and mixing within the building. Mathematical models for vapor intrusion all have incorporated one or more of the formulations and the general applied mathematical formulation of each of these components are described below.

10.4.1.1 Phase Partitioning

Phase partitioning calculations are typically performed to estimate the concentration of a particular contaminant in soil gas from its concentration in any other phase (i.e., groundwater, bulk soil, or non-aqueous phase liquid).

Groundwater to Soil Gas

The concentration in soil gas overlying contaminated groundwater can be estimated as follows:

$$C_{sg} = H \times C_{gw} \tag{10.1}$$

where C_{sg} is the vapor concentration in soil gas (μg/L), H Henry constant (–) and C_{gw} the concentration in groundwater (μg/L).

There may be rate-limiting mass transfer from groundwater to soil gas, so this equation will overestimate the soil gas concentration in some cases; however, that would tend to overestimate the potential for health-risks, which is conservative and therefore appropriate for a screening level assessment. It will generally be valuable to confirm soil gas concentrations by direct sampling and analysis if the calculated values pose a potentially unacceptable risk.

NAPL to Soil Gas

The concentration of a contaminant in soil gas in contact with a NAPL (occurring either as a residual phase in soil or as a lens on top of the water table) can be estimated, assuming that the soil air above the NAPL is saturated with the contaminant and related to the maximal contaminant vapor density (Vd) at 283 Kelvin (Fetter 1993):

$$Vd = \frac{V_p \times M}{R \times T} \tag{10.2}$$

Where Vd is the vapor density (g/m^3), V_p the vapor pressure (Pa), M the molecular weight (g/mol), R the gas constant (8.3144 Pa m^3/mol-K) and T the absolute temperature (K). The concentration in soil air is considered equal to the vapor density as given in the equation below:

$$C_{\text{sg}} = Vd \tag{10.3}$$

Where C_{sg} is the concentration in the soil gas (g/m^3) and Vd the vapor density (g/m^3).

Another approach is the use of the Raoult's Law with the component mole fraction (McQuarrie and Simon 1997) and is presented below.

$$C_{vs} = \frac{x_i \times P_v^i \times M \times W_i}{R \times T} \tag{10.4}$$

Where C_{vs} is the vapor concentration at the source (g/m^3), x_i the mole fraction of component i in the hydrocarbon (mol/mol), P_v^i the pure component vapor pressure of component i (atm), M the molecular weight of component i (g/mol), R the universal gas constant (82.1) [cm^3-atm/mol-K) and T the absolute temperature (K).

The mole fraction, x_i, is calculated from:

$$x_i = \left(\frac{C_T}{C_{mix}}\right)\left(\frac{MW_{mix}}{MW}\right) \tag{10.5}$$

where x_i the mole fraction of component i in the hydrocarbon (mol/mol), C_T the total soil concentration of contaminant i (mg/kg$_{\text{dw}}$), C_{mix} the total soil concentration of the mixture (mg/kg$_{\text{dw}}$), MW_{mix} the weighted average molecular weight of the mixture [g/mol] and MW the molecular weight of component i (g/mol).

If the molecular weights of the component, *i*, and the mixture are similar, this roughly translates to the concentration of the component, *i*, over the concentration of total mixture in the source.

The actual concentration may vary depending on several factors. NAPL mixtures often include some compounds that are aerobically degradable, so if oxygen is present, the concentrations in soil gas above the NAPL may be substantially lower than calculated from the partitioning equations. For NAPL that is at or near the water table and there are significant water table fluctuations, soil vapor concentrations tend to increase when the water table is low and decrease when the water table is high. Many compounds of potential concern for vapor intrusion would have concentrations several orders of magnitude above levels of potential concern for vapor intrusion, so where NAPL occurs in the vadose zone in close proximity to buildings, it is a reasonable precaution to proceed with sampling and analysis of soil gas and/or indoor air, and use the phase partitioning calculations only as an initial guide to which of the compounds present are most likely to be significant contributors to the total risk, in order to select appropriate sampling and analytical methods for their characterization.

Bulk Soil Concentration to Soil Gas

The concentration of a contaminant in soil gas in contact with contaminated soil (i.e., a bulk soil sample) can be estimated as follows:

$$C_{sg} = \frac{C_{bs}}{\frac{1}{\mathrm{H}} \times \left[\mathrm{f_{oc}} \times \mathrm{K_{oc}} + \frac{1}{\rho_b} \times (n_w + n_g) \times H \right]} \tag{10.6}$$

where:

C_{sg}: concentration in the soil gas (μg/L)
C_{bs}: concentration in the bulk soil (μg/$\mathrm{kg_{dw}}$)
H: the contaminant-specific Henry's Law constant [(μg/L-vapor)/(μg/L-H_2O)]
$\mathrm{f_{oc}}$: fraction of organic carbon (–)
$\mathrm{K_{oc}}$: organic carbon partition coefficient (mL/g)
ρ_b: soil bulk density (g/mL)
n_w: the volumetric moisture content [L-H_2O/L-soil]
n_g: the volumetric gas content (= $\mathrm{n_T} - \mathrm{n_w}$) [L-vapor/L-soil]

In practice, there is often a poor correlation between measured soil concentrations and measured soil gas concentrations, so the partitioning calculations should be used with caution. The United States Environmental Protection Agency (US EPA) conducted a study that showed soil sampling and analysis of volatile contaminants by SW846 Method 8240 may have negative biases of a factor of 10 to 1,000 because of loss of volatiles (EPA 1993a). However, others have observed a tendency for partitioning calculations to overestimate the soil vapor concentrations (Provoost et al. 2009). The fraction of organic carbon in some soils is highly variable, and the bulk soil concentration can be significantly higher or lower if there is more or less organic matter present, respectively. Research is required to improve the understanding of this topic.

10.4.1.2 Transport Through a Porous Media

Vapor transport through water-unsaturated soils is typically dominated by diffusion through the vapor phase, because for most VOCs the gas phase diffusion coefficient is relatively high (on the order of 10,000 times higher than aqueous diffusion coefficients).

The vapor mass flux is calculated by Fick's law, as follows:

$$J = -D_{eff} \frac{\partial C}{\partial z} \tag{10.7}$$

where:

J is the vapor mass flux [M/L^2-T]
C is the vapor concentration [M/L^3]
z is the distance over which the concentration change is measured [L]
∂C/∂z is the concentration gradient [M/L^4],
D$_{\mathrm{eff}}$ is the effective diffusion coefficient for the medium, [L^2/T]

In porous media, the effective diffusion coefficient depends on the porosity and water-filled porosity of the medium, as formulated by Millington and Quirk (1961).

$$D_{eff} = D_a \frac{\theta_a^{10/3}}{\theta_T^2} + \frac{D_w}{H} \frac{\theta_w^{10/3}}{\theta_T^2} \tag{10.8}$$

where:

D_a is the free-air diffusion coefficient [L^2/T],
D_w is the aqueous diffusion coefficient [L^2/T],
θ_a is the soil air filled porosity [volume vapor/total volume],
θ_T is the soil total porosity [volume pores/ total volume],
θ_w is the soil water-filled porosity [volume water/total volume], and
H is the dimensionless Henry's Law Constant [molar concentration in gas/ molar concentration in water].

10.4.1.3 Vapor Intrusion into Buildings

Calculating the flow rate for soil gas into buildings is challenging, because it depends on the building pressure or vacuum relative to the pressure in the subsurface beneath the building, the permeability of the soil and fill materials beneath the building, and the permeability of the foundation, all of which vary from building to building and may also vary over time.

A conservative estimate of the potential for vapor intrusion can be developed by assigning a value for the volumetric flow rate of soil gas into a building that is sufficiently high to allow unrestricted entry of vapors diffusing upward to the subfloor region from a deeper source. For a typical residence, this is on the order of 1 to 10 L/min (Johnson 2002).

10.4.1.4 Attenuation Factors

Due to the complexity of the vapor intrusion pathway, some "models" of the pathway use an attenuation factor (AF) approach, rather than the formulations described above, to relate indoor air concentrations to subsurface source concentrations. These approaches are expressed as follows:

$$C\text{air} = C\text{source} \times \text{AF}$$

Where Cair is the indoor air concentration, Csource is the concentration in the soil gas at the source (or some specified depth in the subsurface), and AF is the attenuation factor, which is a value less than one as defined here that accounts for dilution in the building ventilation, degradation in the subsurface, resistance to mass transport imposed by barriers, and similar mechanisms.

The attenuation factors can be estimated at least three different ways. They can be:

1. Derived from empirical data (e.g., Johnson et al. 2002; Mc Donald and Wertz 2007; US EPA 2009);
2. Calculated using mathematical models (described in Section 10.4.2), or
3. Derived using tracer chemicals or marker compounds (Section 10.5.2.4).

10.4.2 Available Vapor Intrusion Models

Mathematical models for vapor intrusion can be used to estimate indoor air concentrations from subsurface vapor concentrations at some depth, perform sensitivity analyses to assess probable ranges of conditions and identify important processes, or provide an expectation of vapor behavior to guide strategies for site characterization or mitigation. If the indoor air concentration can be calculated with reasonable accuracy (i.e., within approximately an order of magnitude), it may not be necessary to collect indoor air samples, particularly where the predicted indoor air concentrations are far below risk-based standards. Models are particularly useful for determining potential indoor air concentrations in situations where a building is not present but may be constructed in the future. Model results may also provide a line of evidence that makes it easier to identify the contribution of vapors from sources inside the building.

Different software models of vapor intrusion are available (commercially or in the public domain). The models most frequently used in Europe and the US are discussed below. The discussion is not intended to express any partiality. It is up to the user to select a model, verify that the formulation and assumptions are consistent with the conditions at a particular site before use, and understand the model capabilities, limitations and sensitivities to various input parameters prior to use. Also, it is generally acknowledged that a qualified and experienced model user improves the proper use of vapor intrusion models.

At present, there is considerable debate about the accuracy of vapor intrusion models. The number of comparisons between carefully applied models and carefully collected field data is limited (Abreu and Johnson 2005). It appears that in most applications, models can provide at best an estimate of indoor air concentrations within an order of magnitude of measured values (Bradley, Patterson and Davis 2009). However, measured indoor air concentrations using 24-h samples can also show up to about one order of magnitude variability (Kuehster et al. 2004), so in this context, the model accuracy is often reasonable.

The available models are subdivided (see Table 10.1) into models that use a mass flux calculation to estimate indoor air concentration and models that use an attenuation factor approach. The models based on mass flux calculations are further divided into those that include diffusive transport of soil air and those that include both diffusive and convective transport. Parameters that most strongly influence the predicted indoor air concentration for all models are source concentration and depth, soil moisture content, building ventilation rate, and oxygen supply (where aerobic degradation is incorporated). For the dilution factor model from Norway and the Johnson and Ettinger (1991) Model (J&E Model), the contribution of the soil parameters is higher than for most of the other models mentioned in this chapter. The exception is the Flemish Vlier-Humaan model, for which the building and physical-chemical parameters equally influence to the predicted indoor air concentration, while variation in the soil parameters does not result in much variation in the indoor air concentration. Because of its appearance in regulation and guidance in the US, much attention has been focused on testing the J&E Model.

Table 10.1 Overview of mathematical models

Item	CSOIL	Vlier-Humaan	JEM	VOLASOIL	RISC	DF Sweden	DF Norway
Compartment/floor							
– slab-on-grade		•	•	•	•		
– concrete basement		•	•	•	•		
– crawl space	•	•		•		•	•
Transport							
– diffusive	•	•	•	•	•		
– diffusive plus convective			•	•	•		
– attenuation factor/empirical						•	•
Source							
– groundwater		•	•	•	•	•	•
– vadose zone	•	•	•	•	•	•	•
– soil gas				•			
Application							
– site specific assessments		•	•	•	•		
– derivation of screening levels	•	•	•			•	•

JEM Johnson and Ettinger model; *DF* Dilution Factor; *X* applies for model

10.4.2.1 Diffusion Models

The CSOIL exposure model was developed in the Netherlands by the National Institute for Public Health and the Environment (RIVM) and used to derive Soil Quality Standards (Rikken et al. 2001). A commercial version is available as the Risk-Human model from the Van Hall Business Centre, which allows the user to perform site-specific Risk Assessments. The model includes only diffusive vapor transport and originates from the HESP model (Van den Berg 1991, 1994, 1995). The model calculates the indoor air concentration for a typical Dutch dwelling with a crawl space as a result of vapor intrusion from groundwater or vadose zone contamination. More information about the commercial version can be found via the website: http://www.risc-site.nl/.

Huijsmans and Wezenbeek (1995) evaluated the accuracy of the CSOIL model and demonstrated that in some situations the functionality of the CSOIL model was not sufficient, e.g., if a concrete basement, slab-on-grade or pure product is present. The indoor air concentration predictions for some of the aromatic VOCs (degradable) were a factor 36 to 360,000 higher than the measured indoor air concentrations. The predictions for some chlorinated VOCs (much less or not degradable) were a factor 2 to 690 higher than the measured indoor air concentrations.

The Vlier-Humaan model is used in Flanders to derive soil clean-up values and to perform site-specific Human Health Risk Assessments. The model calculates a dose and an indoor air concentration as a result of the presence of contaminants in soil or groundwater and compares the dose and indoor air concentration with a TDI (Tolerable Daily Intake) and a TCA (Tolerable Concentration in Air), respectively. The model includes only diffusive vapor transport and originates from the human exposure to soil contaminants (HESP) model (ECETOC 1992). The basic model for volatilization was derived on the basis of publications from Jury (Jury et al. 1983 and 1990) and adapted for certain exposure characteristics that fit the situation in Flanders. Buildings foundations that can be considered include slab-on-grade floor, concrete basement and crawl space. All background information is available in the technical guidance document part 1 to 4 (Provoost and Cornelis 2004a, b, c, d). More information about the commercial version of this model can be found via the website: http://www.risc-site.nl/.

10.4.2.2 Diffusion and Convection Models

The J&E Model is a screening-level model that considers 1-dimensional upward diffusion from a subsurface source through the water-unsaturated zone, advective flow into the building through a foundation crack due to under-pressurization, and dilution in the building due to ventilation. It is in widespread use, since the US EPA commissioned the production of spreadsheet versions of the J&E Model (EQM 2004). The first generation of the J&E Model (Johnson and Ettinger 1991) did not consider biodegradation (which was added in Johnson et al. 1999), barometric pumping (e.g., Massmann and Farrier 1992; Parker 2003), preferential pathways, fractured subsurface media, or other processes that may be important in some circumstances. Therefore, it may not be applicable in specific circumstances, without modifications. Johnson (2002) provides a detailed discussion of the critical input parameters for the J&E Model and guidelines for the model's use.

The National Institute for Public Health and the Environment (RIVM) in the Netherlands developed the VOLASOIL model in the mid-1990s (Waitz et al. 1996) and used it for site-specific indoor air concentration predictions in a decision framework for determining the urgency of remediation. The model includes both diffusive transport and convective transport. Site-specific data can be used as input data.

The VOLASOIL model was reviewed, further verified and extended by Bakker et al. (2006) and Van Wijnen and Lijzen (2006) and now includes three foundation types (i.e., crawl space, concrete basement and slab-on-grade floor). More information about the extended, commercial version can be found via the website: http://www.risc-site.nl/. Van Wijnen and Lijzen (2006) used this model to compare predicted and measured indoor air concentrations due to vapor intrusion from contaminated groundwater into buildings with crawlspaces. The study showed that tetrachloroethylene (PCE) concentrations in air were positively correlated to the groundwater concentration for all compartments (soil, crawl space and indoor air) and that the relationships are significant at the 95% confidence level. Indoor air concentrations were well predicted, with more than 50% of the observations within one order-of-magnitude. However, the observed concentrations for biodegraded products, like vinyl chloride (VC) and cis 1,2-dichloroethene (DCE), were often much lower than the predictions. The observed TCE concentrations in the soil air were in general within one order-of-magnitude higher and lower, though some of the observed concentrations were up to 2 orders-of-magnitude lower than the predictions. Other scenarios (concrete basement and slab-on-grade floor) need to be further evaluated.

Risc is a model for performing fate and transport modelling and Human Health Risk Assessments for contaminated sites. Risc contains various soil vapor emission models for predicting indoor and outdoor air concentrations. Risc can model volatilization from groundwater and the vadose zone to indoor and outdoor air and includes diffusive and convective transport of soil air to the building (BP 2001). The model for volatilization from groundwater to indoor air is derived from ASTM (2004) and for soil to indoor air from Johnson and Ettinger (1991). The model includes consideration of a concrete basement or slab-on-grade floor, but not a crawl space. Risc can be used to estimate the potential for adverse human health impacts (both carcinogenic and non-carcinogenic) and includes nine exposure pathways. More information about the commercial version of Risc can be found on the website http://www.bprisc.com.

An extensive evaluation of vapor intrusion screening models was done by Evans et al. (2002), who evaluated nine vapor intrusions models (including the Risc, J&E Model and VOLASOIL models) for the UK Environment Agency in terms of their contents, benefits, limitations and suitability for modelling vapor intrusion for United Kingdom conditions. According to Evans et al. (2002) no single model was considered to satisfy all the criteria, although it was concluded that the Risc model most closely satisfied the criteria for inclusion in the British contaminated land Exposure Assessment algorithm (CLEA) model. Evans et al. (2002) concluded that the Risc model was relatively easy to use and appeared to be mathematically correct and robust. In addition, the model is also typically conservative in its predictions, which makes it a suitable model for first tier Risk Assessments. The study stated that screening models that calculate indoor air concentrations as a result of groundwater contamination need further verification (Evans et al. 2002).

Several other models have been developed for assessing subsurface vapor transport and/or vapor intrusion to indoor air. Krylov and Ferguson (1998) created a model that is designed for buildings with suspended timber floors or crawlspaces. Lowell and Eklund (2004) used mathematical modeling to assess the distance from a source within which vapor intrusion could occur at levels of concern.

10.4.2.3 Dilution Factor Models

In Sweden, a multi-media exposure model for contaminated sites is used within a decision framework for Risk Assessment and remediation of contaminated areas (Naturvårdsverket 1996). The Swedish Environmental Protection Agency has developed a spreadsheet-based model for the calculation of guideline values and to perform site specific Risk Assessments (Naturvårdsverket 2005). Soil and chemical-specific properties are used to estimate the soil air concentration. The model calculates the indoor air concentration from the soil air concentration via a standard dilution factor (DF) of 1:20,000. The DF is empirically derived from measured data. A standard value is applied for volume and ventilation rate of the house, which is assumed to be a one-compartment house.

In Norway, contaminated sites are managed according to preliminary guidelines published by the Norwegian Pollution Control Authority (SFT). The guidelines were published in 1999, and incorporated a three-tiered approach (SFT 1999) to assess risk at contaminated sites in relation to land use and receptor (child and adult). The generic assumption used for vapor intrusion is a DF between soil gas and indoor air of 1:2,000. Soil and chemical specific properties are used to estimate the soil gas concentration. A standard value is applied for volume and ventilation rate of the one-compartment house (SFT 1995).

The US EPA recently conducted a very detailed analysis of empirical data (measured concentrations of VOCs in groundwater, soil gas, sub-slab-soil gas and indoor air), and performed a detailed analysis to establish the order statistics for the AF (inverse of DF) values for each media to indoor air (US EPA 2008). The sub-slab to indoor air DF values compared favourably to information from the study of radon dilution factors (Little et al. 1992), with an average value of about 1:500, and an upper estimate of about 1:50. Groundwater to indoor air DF values were about an order of magnitude greater, with a slightly larger range, which is to be expected considering the range of geologic materials, water table depths, and other factors involved.

10.4.2.4 Numerical Models

A three-dimensional (3-D) numerical model for vapor intrusion was recently developed at Arizona State University (Abreu and Johnson 2005, 2006), which is the most comprehensive model for vapor intrusion available to date. This model has been used to simulate a number of common scenarios to improve the general understanding of the processes affecting subsurface vapor transport mechanisms (Abreu in press; Abreu et al. 2008). Pennell et al. (2008) and Bozkurt et al. (2008) also have developed a 3-D model using a commercially available generic numerical code and have used the model to investigate the influence of heterogeneous soils on vapor intrusion. 3-D models allow simulations of much more complex geometries than one-dimensional models, and therefore provide additional insight into the expected behavior of vapors under a wider range of conditions.

Subsurface vapor transport (without explicit simulations of buildings) can also be simulated with Air3D (Joss and Baehr 1995), VapourT (Mendoza and Frind 1990a, b), CompFlow Bio (Yu et al. 2009) and a multiphase model by Sleep and Sykes (1989). These models were designed to simulate subsurface contaminant transport through multi-phase or vapor transport, and are flexible enough to simulate heterogeneous soils, complex geometry, advection and diffusion, and other mechanisms that may help to elucidate the important mechanisms involved in vapor intrusion. These models are more complicated and require considerable effort to use, so they have had limited application to vapor intrusion studies to date.

10.5 Sampling and Analysis

The sampling and analysis necessary for site-specific assessment of subsurface vapor intrusion is challenging, because of the low quality standards, multiple alternative sources of vapors at these low concentration levels, and the potential for sampling bias and variability. This chapter will discuss these challenges in detail, and discuss approaches for managing them.

10.5.1 Sampling and Analysis Challenges

Indoor air sampling and analysis may initially seem like the first choice for assessing subsurface vapor intrusion to indoor air. However, the results are almost always difficult to interpret because of background contributions from consumer products, building materials, and even outdoor air sources. Several contaminants have indoor air quality standards that are lower than typical reporting limits for conventional laboratory methods of analysis, in which case, indoor air quality monitoring may not be a viable line of evidence and other lines of evidence will be required. Indoor air quality can also be variable in time and space, depending on wind, barometric pressure, occupant's activities and heating or air conditioning operations, and outdoor concentrations. Therefore, indoor air quality measurement is generally not simple or unambiguous, and sampling should include ambient as well as indoor air.

Soil gas sampling and analysis faces different challenges, particularly related to sample representativeness. When vapor intrusion is a bona fide concern, soil vapor concentrations are typically higher than indoor air concentrations by a factor that accounts for the amount of attenuation in concentrations that occurs as subsurface vapors migrate to indoor air (generally at least a factor of about 100). Therefore, it is generally easier to resolve soil gas concentrations of concern against background concentrations and analytical detection limits.

Groundwater sampling and analysis methods are generally more mature than soil gas sampling methods. However, groundwater is furthest removed from the receptor, and the estimation of indoor air concentrations from groundwater data therefore involves a substantial additional uncertainty.

It is important to design the site-specific assessment with consideration of the challenges facing each sampling and analytical approach. In many cases, multiple lines of evidence may be advisable to avoid potential biases inherent in any single method. This section describes the theoretical considerations, which will help the practitioner in study design.

10.5.2 Pros and Cons of Sampling for Various Soil Compartments

Samples for contaminant analysis may be collected from several different compartments during the course of a vapor intrusion pathway evaluation, including indoor air, soil gas (sub-slab or in native material), and groundwater. Samples from these different compartments have their own particular uses, benefits, and cautions as summarized in Table 10.2. More detailed descriptions of the benefits and limitations of the various sample collection options are included in this section.

10.5.2.1 Shallow Groundwater

Groundwater sampling data are likely to be available at many sites, even where soil gas data have not yet been collected, so it is often the first available line of evidence for screening the vapor intrusion pathway. Groundwater sampling protocols have been developed and refined over the course of the past few decades to the point where many historic causes of bias and variability have been resolved, so the data quality tends to be good. Even where groundwater data are not already available, groundwater samples can often be collected as quickly and as easily as soil gas samples. To the extent practicable, groundwater samples should be collected over a narrow interval (a few ft/1 m or less) (Ter Meer et al. 1999) just below the groundwater table when the data are to be used for assessing the potential for vapor intrusion. Interpretation of groundwater data should consider that groundwater typically is the farthest removed of all compartments from the receptor, and for this reason, should generally be considered a supporting line of evidence. However, for

Table 10.2 Summery of main benefits and limitations of the various sample collection options

Compartment	General description/comment	Benefits	Limitations
Groundwater	Groundwater monitoring wells are installed outside the building footprint. Screened interval must cross the water table, and preferably be short.	Is not intrusive into residence. Simpler access agreements Sampling protocols are more mature than soil gas equivalents	May not be representative of conditions immediately beneath building Existing wells may not have shallow and short screens Requires extrapolation or modeling to estimate indoor air concentrations
Sub-slab soil gas	Probes are drilled through building foundation to collect soil gas sample.	More representative of composition of subsurface vapors that may migrate to indoor air. Impacts of background sources of contaminants are less significant	Intrusive. Requires access to living area, and may cause minor damage to flooring Multiple samples (in time and position) are necessary to estimate representative concentrations beneath entire floor slab Requires extrapolation or modeling to estimate indoor air concentrations
Soil gas	Probes are installed outside of building footprint. Typical depths range from 5 to 30 ft/10 m (but can be deeper).	Is not intrusive into residence. Simpler access agreements Impacts of background sources of contaminants are typically not significant, although equipment contamination may be significant Soil gas program may be used to quantify the significance of biodegradation.	May not be representative of conditions immediately beneath building because of rain-shadow beneath building, and possible impedance to O_2 recharge under floor slab. Requires extrapolation or modeling to estimate indoor air concentrations
Indoor air	Typically, 8 or 24 h sample collected from basement or living area of building	Evaluates the concentration of contaminants of concern at the point of exposure.	Intrusive. Requires access to living area Difficult to distinguish background impacts on analytical results. The measured concentration may not be due to vapor intrusion pathway. Multiple samples (in time and position) are necessary to estimate representative concentration for a 30 year exposure scenario (or longer-duration samples)

a building with a basement depth similar to the depth to the water table ("wet basements"), groundwater may be the only compartment beneath the building that can practically be sampled.

In some cases, groundwater data may not be appropriate for screening the vapor intrusion pathway. For example, if a building is overlying soils contaminated with residuals from historic releases, groundwater quality data may not be relevant at all. If contaminated groundwater is overlain by a fresh water lens, as shown previously in Fig. 10.5, and monitoring well screened intervals are too long to reflect conditions at the water table, the groundwater data may provide misleading results. A fresh water lens of a foot/ 0.3 m or more in thickness may be sufficient to minimize or eliminate volatilization from the groundwater. If volatile contaminants exist in a confined aquifer, the aquitard may act as a partial or substantial barrier to upward migration of contaminants, in which case, vapor intrusion may be negligible. Therefore, at sites where groundwater concentrations appear to be high enough to contribute to vapor intrusion, it may be appropriate to collect confirmatory soil gas data to assess the degree of volatilization from the groundwater table. Transects of shallow groundwater samples can be helpful in some cases for determining an appropriate scope of soil gas and indoor air quality monitoring programs.

For building where the water table is very near the foundation (i.e., the "wet basement" scenario), it may be impossible to collect sub-slab soil gas samples or exterior soil gas samples from adjacent to the building and below the footing, and it may therefore be necessary to rely on shallow groundwater and indoor air data for the assessment. The relative concentrations of various contaminants in each compartment should be similar if the origin of the vapors in indoor air is the subsurface. Contaminants with higher proportions in indoor air relative to other contaminants are likely to either originate from indoor or outdoor air sources not related to the contaminated site, or at least have a contribution from background sources that should be closely reviewed (Weisel et al. 2008).

10.5.2.2 Sub-Slab Soil Gas

Sub-slab soil gas is the gas that exists immediately beneath the floor of the occupied structure, regardless of whether the structure is a slab-on-grade or basement design. Sub-slab soil gas sampling has a number of advantages over sampling of other compartment. It is more representative of composition of subsurface vapors that may migrate to indoor air. It is less susceptible to impacts of background sources of contaminants, although with barometric pressure fluctuations, it is possible for indoor air sources to cause vapors to move from the building. Because soil gas concentrations tend to be high, they are typically easier to resolve against laboratory reporting limits than indoor air samples.

Sub-slab soil gas sampling is relatively simple and can be accomplished with an electric hammer-drill, avoiding the need for a more-costly drilling rig. However, sub-slab sampling has certain drawbacks. It requires an access agreement from the building owner, and is intrusive to the extent that equipment must be brought into the building, dust is generated and floor-coverings may be damaged, all of which is unpleasant for property owners. Relatively little information is available to demonstrate how sub-slab soil gas concentrations vary over time, or in response to barometric pressure changes (Luo et al. 2006). Recent publications have shown that soil gas measurements showed a high degree of spatial variability and expressed the need for a better time resolution (Eklund and Simon 2007; Luo et al. 2006). Tillman and Weaver (2007a) investigated the effect of temporal moisture content variability beneath and external to a building and its effect on vapor intrusion assessment. Some regulatory guidance documents (e.g., US EPA 2006) recommend collecting three sub-slab samples for a building the size of a typical domestic residence in the US to account for spatial variability, so sub-slab sampling efforts are not insignificant, especially if the vapor intrusion assessment includes a neighborhood of residences. Sub-slab soil gas sampling methods may not be practical for buildings with suspended floors and crawlspaces.

10.5.2.3 Soil Gas Samples Collected Adjacent to a Building

Soil gas samples collected beside a building are commonly used to assess vapor intrusion, because they can be collected with less disruption to the building occupants. However, experience to date has shown that vapor concentrations in soil gas samples collected beside the building are not strongly correlated to indoor air concentrations or to sub-slab concentrations (US EPA 2008). There are two possible explanations for this observation (spatial variability in soil vapor concentrations and bias or variability imposed by soil gas sampling methods), described in the two paragraphs below. Resolution of the relative contributions of these two issues is a topic of on-going research.

Soil gas surveys have been commonly used for fast and inexpensive site-wide screening to identify possible areas of VOC releases for decades. However, the quality assurance and quality control needed to achieve the data quality objectives for vapor intrusion assessments is much more stringent. Published guidance for soil gas sampling and analysis to address vapor intrusion data quality objectives is sparse at this time. The ASTM Standard Guide for Soil Gas Monitoring in the Vadose Zone (ASTM D-5314-92) was originally written in 1992, and re-authorized in 2001, since which time there has been a rapid increase in interest in vapor intrusion. The California Department of Toxics Substance Control (CalDTSC) and the Los Angeles Regional Water Quality Control Board (LARWQCB) provides a soil gas sampling advisory (CalDTSC & LARWQCB 2003), although this is being updated at the time of this document's publication. Some non-government organizations have developed good references for soil gas sampling specific for assessing vapor intrusion that provide good information, but are not entirely consistent (e.g., API 2005; EPRI 2005; Geoprobe 2006; ISO 2004; VDI 1998). However, several regulatory guidance documents for vapor intrusion assessment provide no substantial discussion of soil gas sampling methods (e.g., MADEP 1994, 2002; US EPA 2002).

Johnson and Abreu (2003) showed that the depth of the soil gas sample is very important. Figure 10.8 shows simulated normalized (i.e., C/Cmax) concentrations

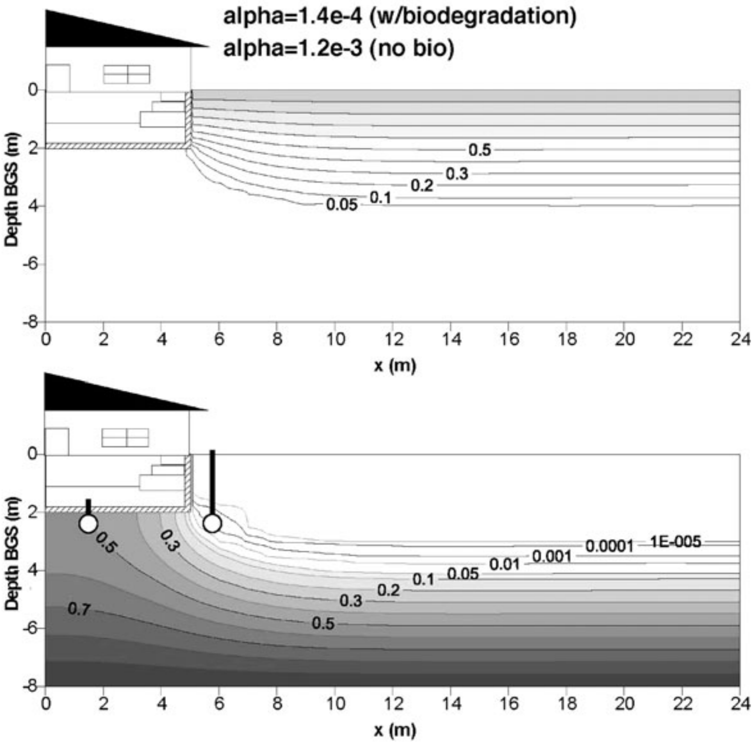


Fig. 10.8 Simulated normalized (i.e., C/Cmax) oxygen (*top*) and hydrocarbon (*bottom*) vapor concentrations and comparison between near-slab and sub-slab sampling (Johnson and Abreu 2003)

of O_2 (upper image) and a typical hydrocarbon (lower image). Based on these simulations, a soil gas sample collected outside the building at a depth just slightly below the foundation could have a concentration a few orders of magnitude lower than the concentrations beneath the building (these two locations are shown as open circles with solid lines above them in the lower image of Fig. 10.8). Soil gas probes should generally be installed to depths 1–2 m (3–6 ft) below the foundation to avoid a false negative bias in the screening (failing to identify a potential vapor intrusion condition), although this may depend on the site-specific soil conditions, depth to the groundwater table, and whether the contaminants of concern are aerobically degradable.

Field sampling and analysis to verify the model simulations is ongoing. EPA recently completed a sampling program to assess horizontal and vertical profiles of vapor concentrations adjacent to a concrete slab, and found significantly lower concentrations beside the slab than those beneath the slab (EPA 2009).

Spatial variability may also need to be considered and some practitioners have suggested that soil gas samples should be collected from adjacent to all four walls of a residence, for example. In general, sample spacing should be proportional to the size of the contaminated area (i.e., large plumes may be adequately characterized with larger spacing between samples than smaller plumes). At some contaminated sites, VOCs in groundwater have been mapped over distances of up to several miles/kilometers, in which case subsurface concentrations (groundwater or soil gas) are unlikely to vary significantly over distances of 50 ft/15 m in the direction of groundwater flow. Concentrations generally change more rapidly in the direction perpendicular to groundwater flow, because transverse dispersion is much less significant than longitudinal dispersion. Sample spacing ultimately must be a site-specific consideration. A comprehensive discussion of vertical profiles and transects of soil gas data is provided by API (2005), which may help guide the selection of an appropriate scope of soil gas data collection.

Temporal variability in deep soil gas samples tends to be low, so multiple samples over time may not be necessary. However, shallower samples can show temporal variability associated with seasonal changes in groundwater levels and infiltration of recharge (McAlary 2008).

10.5.2.4 Indoor Air

Indoor air typically contains dozens of contaminants at detectable concentrations, some of which may be the same contaminants that are present in the subsurface at a particular site, and some of which may be released by indoor sources or present in ambient outdoor air at levels above risk-based target levels. Indoor air sampling will identify detectable contaminants from all three sources (subsurface, indoor and outdoor), and the effort required to resolve the relative contribution from each is usually not trivial. Access agreements will often be required for indoor air sampling, along with a community relations plan, and may stimulate legal, news-media or political issues. Therefore, it is usually preferable to assess subsurface concentrations (near-slab or sub-slab) first, and assess potential indoor air concentrations through empirical or modeled attenuation factors. If the subsurface concentrations are too low to pose a potential risk from subsurface vapor intrusion (and are adequately represented by the available samples), it may be possible to avoid indoor air sampling and the complexities associated with resolving background contributions. Conversely, if the subsurface concentrations are very high, it may be appropriate to proceed with implementing exposure controls or mitigation systems. In areas where subsurface concentrations are neither too low to pose a potential risk, nor high enough to justify preemptive mitigation, indoor air sampling is usually appropriate.

When collecting indoor air samples to evaluate the potential for subsurface vapor migration to indoor air, it is important to consider the contribution of background (indoor) sources and ambient (outdoor) sources, not related to the contaminated site, to indoor air quality. Background or ambient sources may contribute vapors to indoor air at detectable concentrations, or possible concentrations above target levels. There are many potential indoor sources of background contaminants, including: household activities (smoking, cleaning, hobbies); consumer products (gasoline, heating oil, cleaning supplies, glues); and building materials (carpets, paints, glues). Specific contaminants found in household products may be identified through the National Institute of Health Database (http://hpd.nlm.nih.gov/index.htm).

Many of the contaminants considered in the vapor intrusion pathway evaluation are also present in outdoor air (Berry-Spark et al. 2004; McDonald and Wertz 2007), particularly in urban areas. Potential sources of outdoor air impacts include automobile emissions, manufacturing sites, and locations with significant contaminant use (e.g., dry cleaners). In some cases, these outdoor air concentrations may be greater than the target risk-based indoor air concentrations. Outdoor air quality may contribute vapors to indoor air at concentrations above target levels. Therefore it is advisable to include collection and analysis of outdoor air samples in the scope of work conducted for any indoor air quality survey. It may also be informative to collect outdoor air samples when conducting a soil gas survey in urban areas, because ambient air can lead to detectable concentrations of VOCs in soil gas when detection limits are as low as those required to meet soil gas target levels or screening levels.

In the US, several studies of background concentration of contaminants in indoor air have recently been compiled (Dawson 2008; Dawson and McAlary 2008). This compilation identified several contaminants (benzene, tetrachloroethene, chloroform, carbon tetrachloride, and methyl tert-butyl ether) whose average concentrations are similar to the range of common risk-based quality standards. Therefore, there will very likely be issues to resolve regarding the relative contribution of subsurface and indoor sources for vapor intrusion investigations at sites where any of these contaminants are present in the subsurface. Occupant's habits and commercial products vary considerably from house-to-house, so control properties may not provide relevant background data. Therefore, it is generally not recommended to include indoor air quality monitoring in homes outside of the study area in order to assess background concentrations. Alternate approaches have been used to assess indoor air background concentrations. These alternate approaches include:

- Qualitative comparison to published data for background (indoor) and ambient (outdoor) air quality.
- Use of a tracer compounds to select homes with no (or insignificant) subsurface impacts. Tracer compounds are compounds that are present in the subsurface and indoor air at concentrations clearly higher than background levels. This may include compounds that are not common in consumer products and building materials (e.g., 1,1-dichloroethene, cis-1,2-dichloroethene), naturally occurring radon, or other contaminants.
- Evaluation of indoor air concentrations prior to and after operation of a sub-slab venting system (e.g., Folkes 2000), or under conditions where the building pressure is manipulated to be greater than and less than soil gas pressures (Berry-Spark et al. 2005; McHugh et al. 2006).
- Evaluation of the ratio of indoor air to subsurface concentrations for a large data set including a wide range of subsurface concentrations (Ettinger 2003; Johnson et al. 2002).

10.5.2.5 Soil Sampling

Soil sampling is generally not recommended for assessing subsurface vapor intrusion to indoor air because there are no published studies that clearly show a unique relationship between measured soil concentrations and measured soil gas concentrations. The poor correlation may be attributable to inconsistent amounts of volatilization losses during soil sample collection (EPA 1993a). The Encore^{TM} Sampling device has been developed to manage the volatilization losses, but the field preservation with methanol results in a solvent peak during analysis that results in elevated detection limits, which may fail to identify the presence of volatile contaminants at concentrations of potential concern for vapor intrusion.

10.5.3 Analytical Methods

The selection of appropriate laboratory analytical methods should be made by developing a list of volatile contaminants known or suspected to be present at a given site in the subsurface at concentrations above regulatory screening levels, and discussion with a laboratory skilled in analysis of air samples. Analytical reporting limits for indoor air samples should be lower than target indoor air concentrations, unless this is technically impracticable. Analytical reporting limits for soil gas samples can be higher, because soil gas concentrations are always attenuated to some degree by building ventilation and other processes. Target detection limits, possible background levels, and sample volumes may all require consideration in the process of selecting the most appropriate sampling and analytical methods.

Vapor samples may be collected in a variety to devices, including rigid Summa^{TM} canisters, flexible Tedlar bags, and sorbent tubes. Summa^{TM} canisters are stainless steel containers with interior surfaces coated with glass. These containers are advantageous because they are a whole air sample, and there is often sufficient volume in the canister to repeat an analysis if there are questions. Holding times up to 30 days are often acceptable.

Tedlar bags are convenient and cost-effective for soil gas and air sampling. However, they have some limitations that are important to consider. In particular, the holding time for Tedlar bags is shorter than for other containers (typically a few days or less), recovery of certain contaminants is poor, and at very low levels there may be detectable concentrations of some VOCs even in brand new Tedlar bags (Hayes et al. 2006). Tedlar bags are preferred containers for field screening, and VOC analysis with a mobile laboratory, and where reporting limits are not extremely low (i.e., >5 parts per billion by volume [ppbv], or about 0.001 μg/L).

Automatic thermal desorption (ATD) tubes or Volatile Organic Sampling Train (VOST) tubes are cylinders packed with adsorptive media through which air is drawn at a measured rate for a measured time. The mass trapped on the tube is determined by laboratory analysis and the concentration is calculated by dividing the mass by the product of the flow rate and time (i.e., volume of air passed through the tube). ATD or VOST tubes are better suited to heavier contaminants, although contaminants as light as vinyl chloride can be trapped, using carefully selected combinations of adsorbent media, of which there are hundreds available. Combinations of adsorbents are often used where there is a mixture of contaminants of interest with a wide range of sorptive properties. Therefore, an experienced analyst is required to select the appropriate media, considering the suite of contaminants to be analyzed, target detection limits and expected concentration ranges. The analyses are destructive, so it is often advisable to sample in duplicate, even if only one sample is analyzed, in order to allow an opportunity for repeat analysis, if needed. There is also a possibility that the pump might fail (e.g., battery failure), in which case the sample would not be representative of the entire planned sampling interval, in which case the duplicate sample could avoid a repeated mobilization of the sampling crew.

10.5.4 Field Screening Considerations

Field screening using portable instruments is an important part of site-specific assessment of subsurface vapor intrusion to indoor air. Field screening may be used for:

- Rapid assessment of potentially explosive conditions, using an explosimeter or landfill gas meter.
- Confirming that soil gas purging yields reproducible soil gas samples, using a flame ionization detector (FID) or photoionization detector (PID).
- Rapid assessment of preferential pathways such as floor drains, sumps, foundation cracks or rooms within a house that might potentially be contributing background vapors using an FID, PID or mobile laboratory.
- Confirming the integrity of samples through the use of tracers (He, SF_6).
- Assessing biodegradation by screening oxygen and carbon dioxide concentrations using a landfill gas meter.

A variety of field instruments may be applied in vapor intrusion investigations, including photoionization detectors (PIDs) and flame ionization detectors (FIDs), landfill gas meters, tracer (e.g., SF_6 and He) detectors, as well as mobile laboratories equipped with GCs. General information about these various instruments is provided below.

10.5.4.1 Photoionization Detectors (PIDs) and Flame Ionization Detectors (FIDs) for VOC Screening

PIDs and FIDs are both capable of detecting VOCs, but some care must be used in the selection of the appropriate instrument for a particular site, depending on the types of contaminants present, the expected concentration ranges, and the presence of any potential interferences. PIDs come with lamps of different power levels (e.g., 10.3 and 11.7 eV), and the lamp must have a power level higher than the ionization potential of the contaminant(s) of concern to be useful. PIDs can be sensitive to water vapor, and since soil gas is generally humid, a water trap is recommended, especially when the air temperature is lower than the ground temperature, and the risk of condensation is increased. FIDs are not sensitive to water vapor, but they require a source of hydrogen gas to fuel the flame, and are somewhat more challenging to operate.

In general, hand-held PIDs provide reproducible total VOC readings within the range of 1 to 10,000 ppmv. FIDs can provide reproducible readings at somewhat lower concentrations, but still generally above 0.1 ppmv. Target indoor air concentrations are generally lower than these instruments are capable of detecting, but they are very useful for soil gas screening, because soil gas quality standards are generally about 100 times higher than indoor air risk-based quality standards.

10.5.4.2 Landfill Gas Meters for Oxygen, Carbon Dioxide and Methane Concentrations

Landfill gas meters are available to quantify oxygen (O_2), carbon dioxide (CO_2) and methane (CH_4) concentrations to low percent levels. Landfill gas meters are useful for field screening of soil gas prior to sample collection to confirm steady readings, much as specific conductance, temperature, turbidity and dissolved oxygen are used for monitoring groundwater purging prior to sample collection. At sites with aerobically degradable contaminants (especially hydrocarbons), it is common to find soil gas with oxygen concentrations that are clearly lower than atmospheric levels, and carbon dioxide concentrations that are clearly elevated above atmospheric levels, which can be very informative for assessing the influence of biodegradation on the potential for vapor intrusion. The detection limits for these instruments is commonly in the low % range (rather than parts per million by volume [ppmv] or ppbv), but this nevertheless provides ample resolution for field screening for these parameters.

10.5.4.3 Hexafluoride and Helium Meters

Helium (He) and sulfur hexafluoride (SF_6) are the most commonly used tracers for air flow and building ventilation testing. Helium is inexpensive, readily available, non-toxic, and easily detected in the range of 0.01 to 100% using portable helium meters, which provides 4 orders of magnitude resolution for tracer testing. Helium is very useful for leak testing, such as where a small volume of helium is sufficient to test seals in sampling trains. SF_6 meters are sensitive to much lower concentrations ($\sim$100 ppbv), so much less tracer gas is required. Therefore, SF_6 is a preferred tracer when testing air flow in large volumes (e.g., entire building ventilation tests, as described by Howard-Reed et al. (2002)).

10.5.4.4 Mobile Laboratories

Mobile laboratories for soil gas surveys have been available in various forms for almost two decades, however, in many cases, the reporting limits were in the range of 1,000 $\mu g/m^3$ (roughly 200 ppbv for many VOCs), which is considerably higher than fixed laboratory reporting limits (e.g., 0.5 ppbv for US EPA Method TO-15). Recent advances in mobile laboratory technology have enabled TO-15 analyses to be performed by mobile laboratories, provided strict QA/QC protocols are followed. Mobile laboratories offer the advantage of real-time information, which can be used to guide the scope of site-assessment activities and in many cases provide adequate site characterization in a single mobilization, instead of multiple phases.

10.6 Mitigation

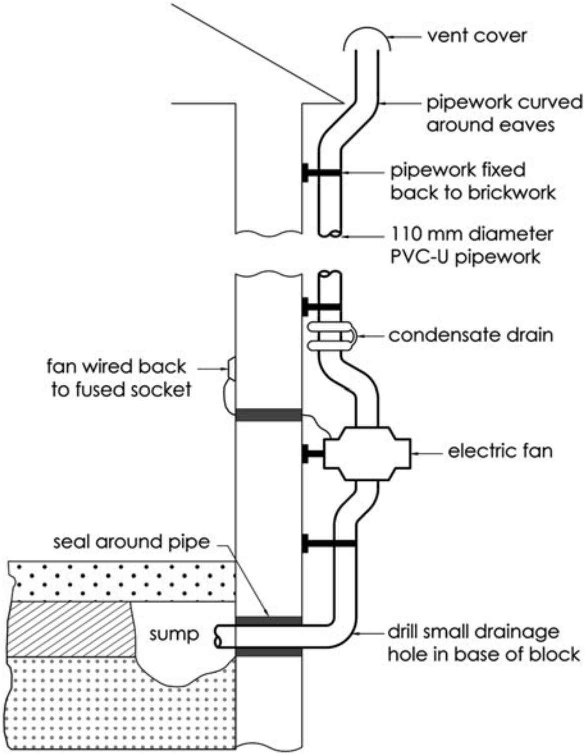
Subsurface vapor mitigation systems may be implemented either because concentrations are high enough to pose a potential risk or as a precaution where Risk Management decisions are intentionally conservative. This section discusses the various options available and the expected effectiveness for each. In practice, mitigation measures that must be added as a retro-fit to an existing building will generally be different than measures which can be integrated into the design and construction of a new building. So this section is presented in two subsections, accordingly.

10.6.1 Methods/Technologies for Existing Buildings

10.6.1.1 Sub-Slab De-Pressurization

Sub-slab depressurization is the most common form of control for subsurface vapor intrusion. In simple terms, a vacuum is applied to the region beneath the floor slab, maintaining a pressure differential sufficient to prevent vapor intrusion. A concrete floor slab is usually underlain by granular fill for structural support, and the granular fill is usually sufficiently permeable to facilitate extraction of sub-slab soil gas and propagation of a vacuum. As a result, a single extraction point connected to a fan of approximately 50–150 Watts is usually sufficient to achieve reductions in indoor air concentrations by 2 or 3 orders of magnitude. Monitoring to confirm the effectiveness usually involves demonstrating that the vacuum is propagating across the floor to some degree (typically 6–9 Pascals, ASTM 2001). These systems are often identical to radon mitigation systems. The fans are rated for continuous duty and last several years between replacements (EPA 1993b). Figure 10.9 shows a typical design. Figure 10.10 shows performance monitoring data demonstrating a reduction of 3 orders of magnitude in concentrations of 11DCE after initiation of sub-slab depressurization (Folkes 2000, 2002). Concentrations of several other contaminants were not affected, which is a clear indication that the source of these other vapors was background contributions. These figures also point out that verification monitoring by indoor air sampling and analysis is problematic, because background

Fig. 10.9 Typical design for a sub-slab venting system (schematic)



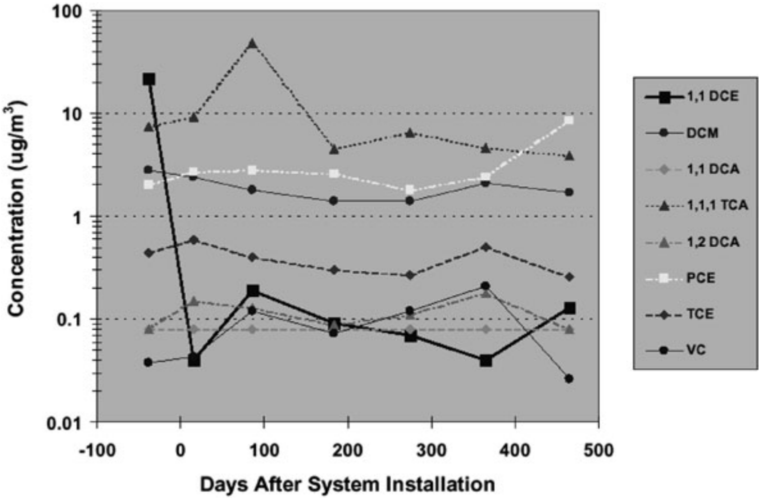


Fig. 10.10 Performance monitoring data from a sub-slab depressurization system (Folkes 2000)

sources cannot be eliminated, therefore, vacuum and flow monitoring are important.

Sub-slab depressurization is a containment technology, and is not designed to remove the source of the subsurface vapors. Therefore, monitoring and maintenance will generally be required to ensure that the systems continue to operate as designed, as long as there is a continuing source of vapors in the subsurface.

10.6.1.2 Soil Vacuum Extraction

In some cases, it may be advantageous to control subsurface vapor intrusion by removing the source of subsurface vapors, rather than, or in addition to blocking the pathway to indoor air. Soil vacuum extraction (SVE) will often be the preferred technology for this approach.

SVE is similar to sub-slab venting, but at a larger scale, with a more aggressive design. SVE is generally intended to remove soil vapors from throughout the vadose zone, accomplishing containment in the short-term and removal of the source of vapors over time. SVE systems will generally require more infrastructure (larger blowers, water knock-out, off-gas treatment through activated carbon filters, and sometimes telemetry and programmable logic controls), permits, regular monitoring, and progress reporting.

10.6.1.3 Building Pressurization

Building pressurization may be as effective as sub-slab de-pressurization for maintaining a pressure gradient across the floor slab sufficient to reduce or eliminate subsurface vapor intrusion. For commercial and industrial buildings where HVAC units are usually mounted on the rooftop and blowing air into the building, it may simply be a matter of running the fans continuously, increasing the air flow rate, and/or installing additional units. In some climates, the energy required to heat or cool air may make this alternative prohibitively expensive. Long-term administrative controls would be required and may be challenging to monitor or audit.

10.6.1.4 Sealing Cracks, Sumps, Sewers, and Other Potential Conduits

Just as sealants can be used to prevent water from entering a basement through discontinuities, they may also prevent or reduce vapor intrusion. Sealants made with current technology (urethane) are very durable and can be expected to last for decades. Holes dilled along cracks can be used to inject sealants to facilitate good penetration and adherence. If an industrial building is planned for re-development, it is often a good idea to seal the entire floor, since historic releases of solvents may have permeated the pores within the concrete. Where cracks are accessible to be sealed, this method may result in a reduction in vapor intrusion, but the magnitude of the reduction achieved by sealing alone will probably be less than the reduction achieved by methods that actively manipulate pressure gradients.

In finished or partially finished basements, it may be difficult to locate and seal cracks, as these may be obscured behind floor and wall coverings. In the short-term after application, some sealants will emit VOC vapors, which should be considered prior to subsequent indoor air sampling and analysis.

10.6.1.5 Air Filtration

Indoor air quality can be improved using filtration units, for example a drum of activated carbon fitted with a blower or fan that circulates indoor air through the carbon, trapping VOCs and SVOCs. These filters typically use activated carbon to trap VOC vapors, and come with an integral fan unit that circulates air through them. They are readily available, can be set up and running in very short order, and provide tangible results almost immediately. The activated carbon requires replacement on a regular basis, which makes this technology generally less attractive for long-term applications.

10.6.2 Methods/Technologies for Future Buildings

Brownfield re-development is a growing opportunity, but vapor intrusion concerns must be considered and in some cases, may pose a potential risk to a planned or hypothetical future building. The technologies described in Section 10.6.1 can be used after building construction, but there are also several methods of mitigation that can be incorporated in the building design or construction, as described below.

10.6.2.1 Intrinsically Safe Building Design

Podium style construction is increasingly common in low-lying areas that are prone to floods, but this design may also be worth considering for redevelopment of land with potential vapor intrusion concerns. Ground-level or underground levels may be used for vehicle parking, and as such, would have adequate ventilation to manage vehicle exhaust, which also would minimize the potential for subsurface vapor intrusion concerns.

10.6.2.2 Vapor Barriers and Ventilation Layers

Subsurface vapor control can be achieved by a vapor barrier and ventilation layer beneath a building. Incorporating these features is relatively inexpensive if placement occurs before the building slab is constructed and the utilities are placed. The most common design consists of a highly permeable layer of granular fill with vent pipes overlain by welded plastic sheets (geomembrane). The vent pipes are connected to a common header that is connected to a blower that can either run continuously, or be automated to run when sensors beneath the building indicate that vapors are accumulating. Construction quality assurance supervision is essential, especially to ensure that the seams of the geomembrane are sealed properly, and that the membrane is not perforated prior to placement of the concrete slab. This can be tested by blowing smoke beneath the barrier and visually inspecting the barrier for any signs of escaping smoke, which should be repaired before the construction proceeds.

Acknowledgments Portions of this chapter were drawn from Reference Handbook for Site-Specific Assessment of Subsurface Vapor Intrusion to Indoor Air by the Electric Power Research Institute (EPRI), document number 1008492, with editing as appropriate to address recent research and practical experience. EPRI's permission to use this material is hereby acknowledged and greatly appreciated. The lead author was the same for the EPRI document and this chapter.

References

Abreu LDV, Johnson PC (2005) Effect of vapor source-building separation and building construction on soil vapor intrusion as studied with a three-dimensional numerical model. Environ Sci Technol 39(12):4550–4561

Abreu LDV, Johnson PC (2006) Simulating the effect of aerobic biodegradation on soil vapor intrusion into buildings: influence of degradation rate, source concentration, and depth. Environ Sci Technol 40(7):2304–2315

Abreu L (in press) Conceptual site models: theoretical overview of vapor intrusion processes. A white paper prepared for the United States environmental protection agency

Abreu L, Ettinger R, McAlary T (2008) Simulated soil vapor intrusion attenuation factors including biodegradation for petroleum hydrocarbons. Accepted in groundwater monitoring and remediation, November 2008

American Petroleum Institute (API) (2005) Collecting and interpreting soil gas samples from the vadose zone: a practical strategy for assessing the subsurface-vapor-to-indoor-air migration pathway at petroleum hydrocarbon sites, publication #4741

API (2006) Downward solute plume migration: assessment, significance, and implications for characterization and monitoring of "diving plumes". API soil and groundwater technical task force bulletin 24, April, 2006

American Society of Heating Refrigerating and Air Conditioning Engineers (ASHRAE) (1985) ASHRAE handbook (1985) fundamentals. Chapter 22. Atlanta, GA

American Society for Testing and Materials (ASTM) (2001) Standard practice for installing radon mitigation systems in existing low-rise residential buildings, designation: E 2121 – 03 Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, US

American Society for Testing and Materials (ASTM) (1992) Standard guide for soil gas monitoring in the vadose zone, ASTM D 5314-92. American society for testing and materials. West Conshohocken, PA, Annual Book of ASTM Standards, v. 04:09

ANSI/ASHRAE (2004) ANSI/ASHRAE Standard 62.1-2004. Ventilation for acceptable indoor air quality and addendum at http://www.ashrae.org/technology/page/132, August 15, 2010

ASTM (2004) Standard guide for risk based corrective action, E2081-00-2004E1, ASTM international

Bakker J, Lijzen JPA, Otte PF, Van Wijnen HJ, Swartjes FA, Baars BJ (2006) Site-specific human-toxicological risk assessment of soil contamination with volatile compounds. National institute of public health and environmental protection, RIVM report 711701049, RIVM, Bilthoven, the Netherlands

Berry-Spark K, Ettinger E, McAlary T, DeFlaun M, Finn J, Lennon J, Sweeney C (2004) Vapor intrusion at a former manufactured gas plant site – assessing background contributions to indoor air. Platform presentation at the Midwestern states risk assessment symposium, Indianapolis, IN, August, 2004

Berry-Spark K, McAlary T, Bice N, Zeeb P (2005) Mitigation of the vapor intrusion pathway at four chlorinated solvent sites. Subsurface vapor intrusion to indoor air: an update. the 14th symposium in GRA's series on groundwater contaminants, May 25, 2005, San Jose, CA

Bozkurt O, Pennell KG, Suuberg EM (2008) Evaluation of the vapor intrusion process for non-homogeneous soils. Ground Water Monitor Remediation. Accepted November 2008

Bradley M, Patterson BM, Davis GB (2009) Quantification of vapor intrusion pathways into a slab-on-ground building under varying environmental conditions. Environ Sci Technol 43(3):650–656. doi:10.1021/es801334x

British Petroleum (BP) (2001) Risc v4.03 – risk assessment model for soil and groundwater applications. User's manual, BP Oil international, Sunbury UK

California Environmental Protection Agency (CalEPA) and Regional Water Quality Control Board (RWQCB) (2003) Advisory – active soil gas investigation. jointly issued by the RWQCB, Los Angeles region and the department of toxic substances control. January 28, 2003. www.dtsc.ca.gov/PublicationsForms/index.html

Davis GB, Rayner JL, Trefry MG, Fisher SJ, Patterson BM (2005) Measurement and modeling of temporal variations in hydrocarbon vapor behavior in a layered soil profile. Vadose zone J, Soil SCI SOC AMER, Madison, USA, SN 1539-1663, year 2005, 4(2):225–239

Dawson H (2008) Overview of indoor air "background" concentrations & overview of vi database, a platform presentation at the EPA Workshop on Vapor Intrusion at the AEHS 18th annual west coast conference on soils, sediment and water, San Diego, March 14, 2008

Dawson H, McAlary T (2008) A compilation of order statistics for VOCs from post-1990 indoor air concentration studies in North American residences unaffected by subsurface vapor intrusion. Accepted in groundwater monitoring and remediation, November, 2008.

DeVaull GE, Ettinger RA, Gustafson JB (2002) Chemical vapor intrusion from soil or groundwater to indoor air: significance of unsaturated zone biodegradation of aromatic hydrocarbons. Soil Sediment Contamin 11(4):625–641

DeVaull GE, Ettinger RA, Salinitro JP, Gustafson JB (1997) Benzene, toluene, ethylbenzene and xylenes (BTEX) degradation in vadose zone soils during vapor transport: first order rate constants. Proceedings of the API/NGWA petroleum hydrocarbons and organic chemicals in groundwater: prevention, detection and remediation conference. Houston, TX. November 12–14, pp 365–379

ECETOC (1992) Hazard assessment of chemicals contaminants in the soil, technical report no. 40 – reviewed appendix 3, ECETOC, Brussels

Eklund BM, Simon MA (2007) Concentration of tetrachloroethylene in indoor air at a former dry cleaner facility as a function of subsurface contamination: a case study. J Air 7 Waste Manage Assoc, Pittsburgh, PA, SN 1047-3289, 57(6):753–760

Eklund B, Folkes D et al (2007) An overview of state approaches to vapor intrusion, environmental manager. Air & waste management association, February 2007

Electrical Power Research Institute (EPRI) (2005) Reference handbook for site-specific assessment of subsurface vapor intrusion to indoor air. Electric power research institute, Palo Alto, CA. http://mydocs.epri.com/docs/public/000000000001008492.pdf, August 15, 2010

Environmental Quality Management, Inc. (EQM) (2004) User's guide for evaluating subsurface vapor intrusion into buildings. Prepared by EQM, Inc. for U.S. EPA. http://www.epa.gov/oswer/riskassessment/airmodel/pdf/guide.pdf, October 13, 2009

Evans D, Herbs I, Wolters RM, Boddington RTB, Hall DH (2002) Vapour transport of soil contaminants. Environment agency, R&D technical report P5-018/TR

Ettinger RA (2003) The impact of background concentrations on vapor intrusion assessment. Invited platform presentation at the California groundwater resources association symposium "subsurface vapor intrusion to indoor air: when is soil and groundwater contamination and indoor air issue?"

Fetter CW (1993) Contaminant hydrology. Macmillan, New York, NY

Fischer JM, Baker RJ, Baehr AL (1996) Determination of vapor-phase diffusion coefficients for unsaturated-zone sediments at Galloway-Township, New Jersey. In Morganwalp DW, Aronson DA (eds) U.S. geological survey toxic substances hydrology program – proceedings of the technical meeting, Colorado Springs, Colorado, September 20–24, 1993, US Geol Surv Water Resour Invest Rep 94-4015

Fitzpatrick NA, Fitzgerald JJ (1996) An evaluation of vapor intrusion into buildings through a study of field data. Presented at the 11th annual conference on contaminated soils, University of Massachusetts at Amherst, October 1996, pp 1–15

Fitzpatrick NA, Fitzgerald JJ (2002) An evaluation of vapor intrusion into buildings through a study of field data. Soil Sediment Contamin 11(4):603–623

Folkes D (2000) Vapor intrusion assessment and mitigation – practical issues and lessons learned. Platform presentation at the RCRA national meeting EI forum, August 16, 2000, Washington, DC. http://www.clu-in.org/eiforum2000/agenda.cfm

Folkes DJ (2002) Design, effectiveness, and reliability of sub-slab depressurization systems for mitigation of chlorinated solvent vapor intrusion, US EPA ORD Seminar on Indoor Air Vapor Intrusion, San Francisco, December 2002

Geoprobe Systems, Inc (2006) Direct-push installation of devices for active soil gas sampling and monitoring. Technical bulletin #3099, May, 2006

Hayes HC, Benton DJ, Khan N (2006) Impact of sampling media on soil gas measurements. Proceedings of A&WMA specialty conference “vapor intrusion – the next great environmental challenge”, January 25–27, 2006, Philadelphia, PA

Hers I, Zapf-Gilje R, Li L, Atwater J (2000) Validation of models used to predict indoor air quality from soil and groundwater contamination. Proceedings of the API/NGWA petroleum hydrocarbons and organic chemicals in groundwater: prevention, detection and remediation conference. Anaheim, CA, November 15–17, p 213

Howard-Reed C, Wallace LA, Ott WR (2002) The effect of opening windows on air change rates in two homes, ISSN 1047-3289. J Air Waste Manage Assoc 52:147–159

Huijsmans KGA, Wezenbeek JM (1995) Validation of the CSoil model intended to quantify human exposure to soil pollution. Contamin Soil 1:621–622

ISO (2004) Soil quality, sampling. Part 7: guidance on sampling of soil gas. Draft international standard ISO/DIS 10381-7

ITRC (2007) Vapor intrusion pathway: a practical guideline. Interstate technology and regulatory council, January 2007

Johnson PC, Ettinger RA (1991) Heuristic model for predicting the intrusion rate of contaminant vapors into buildings. Environ Sci Technol 25:1445–1452

Johnson PC, Kemblowski MW, Johnson RL (1999) Assessing the significance of subsurface contaminant vapor migration to enclosed spaces: site-specific alternatives to generic estimates. J Soil Contamin 8(3):389–421

Johnson PC, Abreu L (2003) Confusion? Delusion? What do we really know about vapor intrusion. Groundwater resources association symposium on groundwater contaminants; Subsurface vapor intrusion to indoor air: when is soil and groundwater contamination an indoor air issue? September 30, 2003 (San Jose) and October 1, 2003 (Long Beach)

Johnson PC, Ettinger RA, Kurtz J, Bryan J, Kester JE (2002) Migration of soil gas vapors to indoor air: determining vapor attenuation factors using a screening-level model and field data from the CDOT-MTL Denver, Colorado site, newsletter from the American Petroleum Institute, number 16

Johnson PC (2002) Sensitivity analysis and identification of critical and non-critical parameters for the Johnson and Ettinger (1991) vapor intrusion model. API soil and groundwater research bulletin no. 17. API. May

Joss CJ, Baehr AL (1995) Documentation of AIR3D, an adaptation of the ground-water-flow code MODFLOW to simulate three-dimensional air flow in the unsaturated zone. U.S. Geol. Survey Open-File Rep. 94-533

Jury WA, Russo D, Streile G, Abd HE (1990) Evaluation of volatilization by organic chemicals residing below the soil surface. Water Resour Res 26(1):13–20

Jury WA, Spencer WF, Farmer WJ (1983) Behavior assessment model for trace organics in Soil. I. Model description. J Environ Qual 12(4)

Krylov VV, Ferguson CC (1998) Contamination of indoor air by toxic soil vapours: the effects of subfloor ventilation and other protective measures. Building Environ 33(6):331–347

Kuehster T, Folkes D, Wannamaker E (2004) Seasonal variation of observed indoor air concentrations due to vapor intrusion at the redfield site, Colorado. Midwestern states risk assessment symposium Indianapolis, August 26

Lahvis MA, Baehr AL (1996) Estimation of rates of aerobic hydrocarbon biodegradation by simulation of gas transport in the unsaturated zone. Water Resour Res 32:2231–2249

Lahvis MA, Baehr AL (1997) Documentation of R-UNSAT, a computer model for the simulation of reactive, multispecies transport in the unsaturated zone, U.S. geological survey open-file report 97-630, 104 pp

Lahvis MA, Baehr AL, Baker RJ (1999) Quantification of aerobic biodegradation and volatilization rates of gasoline hydrocarbons near the water table under natural attenuation conditions. Water Resour Res 35(3):753–765

Little JC, Daisey JM, Nazaroff WW (1992) Transport of subsurface contaminants into buildings: an exposure pathway for volatile organics. Environ Sci Technol 26(11):2058–2066

Lowell PS, Eklund B (2004) VOC emission fluxes as a function of lateral distance from the source. Am Inst Chem Eng: Environ Prog 23(1):52–58. Published online in Wiley InterScience (www.interscience.wiley.com). doi:10.1002/ep.10007

Luo E, Dahlen P, Johnson PC, Peargin T, Creamer T (2009) Spatial variability of soil gas concentrations near and beneath a building overlying shallow petroleum hydrocarbon-impacted soils. Ground Water Monit Remediat 29(1):81–91

Luo H, Dahlen P, Johnson P, Creamer T, Peargin T, Lundegard P, Hartman B, Abreu L, McAlary T (2006) Spatial and temporal variability in hydrocarbon and oxygen concentrations beneath a building above a shallow NAPL source. Platform presentation at the Battelle conference on remediation of chlorinated and recalcitrant compounds, Monterrey CA, May 2006

Massachusetts Department of Environmental Protection (MADEP) (2002) Indoor air sampling and evaluation guide. WSC Policy No. 02-430. Commonwealth of Massachusetts, Boston, MA, April

MADEP (1994) Background documentation for the development of MCP numerical standards. Massachusetts department of environmental protection, USA

Massmann J, Farrier DF (1992) Effects of atmospheric pressures on gas transport in the vadose zone. Water Resour Res 28(3):777–791

McAlary TA (2008) Lessons from petroleum hydrocarbon and chlorinated solvent sites extensively monitored for vapor intrusion. Platform presentation at the EPA workshop on vapor intrusion at the AEHS 18th annual west coast conference on soils, sediment and water, San Diego, March 14, 2008

McAlary T, Nicholson P, Bertrand D, Abreu L, Ettinger R (2007) a case study on the influence of aerobic biodegradation on vapor intrusion at a former refinery property. Platform presentation at the air and waste management association's specialty conference on vapor intrusion, Providence, RI, September 27, 2007

McAlary TA, Berry-Spark K, Krug TA, Uruskyj JM (2004) The fresh water lens and it's effects on groundwater to indoor air attenuation coefficients, a poster presentation at the US EPA vapor intrusion workshop held at the AEHS 14th annual west coast conference on soils, sediment and water, San Diego, March 15–18, 2004

McAlary TA (1989) Comparison of measured and calculated reactive vapor diffusion coefficients in unsaturated sands. M.Sc. Thesis, Department of Earth Sciences, University of Waterloo, Waterloo, ON, Canada

McCarthy KA, Johnson RL (1991) The transport of volatile organic compounds across the capillary fringe. Draft submitted to water resources research, January 9, 1991

McDonald GJ, Wertz WE (2007) PCE, TCE, and TCA vapors in subslab soil gas and indoor air: a case study in upstate New York, Ground water monitoring and remediation, Blackwell, Oxford, England, SN 1069-3629, year 2007, 27(4), pp 86–92

McHugh TE, De Blanc PC, Pokluda RJ (2006) Indoor air as a source of VOC contamination in shallow soils below buildings. Soil Sediment Contamin, Taylor and Francis, Philadelphia, PA, SN 1532-0383, 15(1):103–122

McQuarrie DA, Simon JD (1997) Physical chemistry: a molecular approach, chapter 24, University Science Books, 55D Gate Five Road, Sausalito, CA 94965

Mendoza CA, Frind EO (1990a) Advective-dispersive transport of dense organic vapours in the unsaturated zone 1. Model development. Water Resour Res 26(3):379–387

Mendoza CA, Frind EO (1990b) Advective-dispersive transport of dense organic vapours in the unsaturated zone II. Sensitivity analysis. Water Res Res 26(3):388–398

Millington RJ, Quirk JP (1961) Permeability of porous solids, Trans. Faraday Soc 57:1200–1207

Naturvårdsverket (1996) Development of generic guideline values. Model and data used for generic guideline values for contaminated soils in Sweden. Swedish Environmental Protection Agency, NV Report No. 4639, Stockholm, Sweden

Naturvårdsverket (2005) Beräkningsmodell för riktvärden för mark, remissversion, Swedish Environmental Protection Agency, Stockholm, Sweden

Ostendorf DW, Kampbell DH (1991) Biodegradation of hydrocarbon vapors in the unsaturated zone. Water Resour Res 27(4):453–462

Parker JC (2003) Physical processes affecting the natural depletion of volatile organic compounds in soil and groundwater. Vadose Zone J 2:222–230(2003)

Pennell KG, Bozkurt O, Suuberg EM Development and application of a 3-D model for evaluating site-specific features on vapor intrusion rates in homogenous geologies. J Air Waste Manage Assoc. Accepted September 2008

Provoost J, Cornelis C (2004a) Technical guidance document, part 1 – derivation of soil screening levels and background values, advise on behalf of the Flemish Public Waste Agency (OVAM), report in Dutch, April 2004

Provoost J, Cornelis C (2004b) Technical guidance document, part 2 – Site specific risk assessment for contaminated sites, advise on behalf of the Flemish Public Waste Agency (OVAM), report in Dutch, April 2004

Provoost J, Cornelis C (2004c) Technical guidance document, part 3 – background document with formulas for the multi-media model Vlier-Humaan, advise on behalf of the Flemish Public Waste Agency (OVAM), report in Dutch, 2004

Provoost J, Cornelis C (2004d) Technical guidance document, part 4 – background document for chemical properties and specifications, advise on behalf of the Flemish Public Waste Agency (OVAM), report in Dutch, 2004

Provoost J, Reijnders L, Swartjes F, Bronders J, Seuntjens P, Lijzen J (2008) Functionality and accuracy of seven vapour intrusion screening models for VOC in groundwater. J Soils Sediments – Protection, Risk Assess Remediation, ISSN 1439-0108, doi:10.1007/s11368-008-0036-y

Provoost J, Bosman A, Reijnders L, Bronders J, Touchant K, Swartjes F (2009) Vapour intrusion from the vadose zone – seven algorithms compared. J Soils Sediments – Protection, Risk Assess Remediation, ISSN 1439-0108, doi:10.1007/s11368-009-0127-4

Rikken MGJ, Lijzen JPA, Cornelese AA (2001) Evaluation of model concepts on human exposure. Proposals for updating the most relevant exposure routes of CSOIL. National institute of public health and environmental protection, RIVM report 711701022, RIVM, Bilthoven, the Netherlands

Roggemans S, Bruce CL, Johnson PC, Johnson RL (2001) Vadose zone natural attenuation of hydrocarbon vapors: an empirical assessment of soil gas vertical profile data. API soil and groundwater technical task force bulletin no. 15. American Petroleum Institute. Washington, DC

Roggemans S (1998) Natural attenuation of hydrocarbon vapors in the unsaturated zone, M.S. Thesis, Arizona State University

SFT (1995) Management for contaminated land: preliminary guidelines for executive procedures, report 97:01, Norwegian Pollution Control Authority, Oslo, ISBN 82-7655-030-4

SFT (1999) Management of contaminated land, report 99:06, Norwegian Pollution Control Authority, Oslo, ISBN 82-7655-192-0, 1 January 1999

Sleep BE, Sykes JF (1989) Modeling the transport of volatile organics in variably saturated media. Water Resour Res 25(1):81–92

Ter Meer J, Lubbers R, Sinke A (1999) Risicoreductie van vluchtige verbindingen door afbraak in de onverzadigde zone (Risk reduction for volatile contaminants by degradation in the vadose zone), TNO-report, TNO-MEP-CR99/089, 1999. http://www.bodembreed.nl/inhoud.asp?id=17229

Tillman FD, Weaver JW (2007a) Temporal moisture content variability beneath and external to a building and the potential effects on vapor intrusion risk assessment, Science of the total environment, Elsevier Science Bv, Amsterdam, Netherlands, SN 0048-9697, 379(1):1–15

Tillman FD, Weaver JW (2007b) Parameter sets for upper and lower bounds on soil-to-indoor-air contaminant attenuation predicted by the Johnson and Ettinger vapor intrusion model. Atmos Environ 41(27):5797–5806

United States Environmental Protection Agency (US EPA) (1993a) Behavior and determination of volatile organic compounds in soil: a literature review. Report EPA/600/SR-93/140

United States Environmental Protection Agency (US EPA) (1993b) Radon reduction techniques for existing detached houses, U.S. EPA ORD

United States Environmental Protection Agency (US EPA) (2002) OSWER draft guidance for evaluating the vapor intrusion to indoor air pathway from groundwater and soils (subsurface vapor intrusion guidance). November 29.

United States Environmental Protection Agency (US EPA) (2006) Assessment of vapor intrusion in homes near the Raymark superfund site using basement and sub-slab air samples (DiGiulio DC, Paul CJ), EPA/600/R-05/147

United States Environmental Protection Agency (US EPA) (2008) U.S. EPA's vapor intrusion database: preliminary evaluation of attenuation factors, Draft March, 2008. http://iavi.rti.org/attachments/OtherDocuments/OSWER_Database_Report_Combined_3-4-08.pdf

US EPA (2009) Vertical distribution of VOCs in Soils from Groundwater to the Surface/Subslab EPA/600/R-09/073, August, 2009, 326 pp

VDI (1998) Measurement of organic soil pollutants techniques of active sampling of soil gas. VDI-3865 Blatt 2

Van den Berg R (1991/1994/1995) Human exposure to soil contamination: a qualitative and quantitative analysis towards proposals for human toxicological intervention values (partly revised edition). National institute of public health and environmental protection, RIVM report 755201011, Bilthoven, the Netherlands

Van Wijnen HJ Lijzen JPA (2006) Validation of the VOLASOIL model using air measurements from Dutch contaminated sites – concentrations of four chlorinated compounds. National institute for public health and the environment, RIVM report 711701041, RIVM, Bilthoven, the Netherlands

Waitz MFW, Freijer JI, Kreule P, Swartjes FA (1996) The VOLASOIL risk assessment model based on CSOIL for soils contaminated with volatile compounds. National institute of public health and the environment (RIVM), report no. 715810014

Weisel CP, Alimokhtari S, Sanders PF (2008) Indoor air VOC concentrations in suburban and rural New Jersey. Environ Sci Technol 42(22):8231–8238. doi:10.1021/es8005223

Yu S, Unger AJA, Parker B (2009) Simulating the fate and transport of TCE from groundwater to indoor air. J Contamin Hydrol Jul 21, 107(3–4):140–161. Epub 2009 May 8

Chapter 11
Human Exposure Pathways

Mark Elert, Roseline Bonnard, Celia Jones, Rosalind A. Schoof, and Frank A. Swartjes

Abstract Depending on land use and corresponding human activities, a number of exposure pathways are relevant for human exposure. In this chapter, six important pathways are described, i.e., exposure through consumption of vegetables, consumption of animal products, consumption of domestic water, inhalation of vapours outdoors, inhalation of dust particles (indoors and outdoors) and dermal uptake via soil material (outdoors and indoors). Note that these exposure pathways follow different exposure routes to enter the human body, i.e., oral, inhalation and dermal routes, respectively. Human exposure through all oral and inhalative exposure pathways described in this chapter (so excluding the dermal uptake exposure pathway), follow a similar pattern. This pattern includes three steps. Firstly, the transfer of contaminants from one of the mobile phases of the soil (pore water or soil gas) into a so-called contact medium. Secondly, the intake of that contact medium by human beings. And thirdly, the uptake of part of the contaminants from the contact medium into the blood stream and target organs and the corresponding excretion of the remaining part of the contaminants. For each of the pathways the significance, conceptual model, an example of mathematical equations and of the input parameters is described in this chapter, in detail. Moreover, attention is given to the reliability and limitations of the calculations.

Contents

M. Elert (✉)
Kemakta Konsult, Stockholm, Sweden
e-mail: mark@kemakta.se

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_11,

11.1 Introduction

11.1.1 Relevant Pathways

A number of exposure pathways may be relevant for human exposure, depending on land use and corresponding human activities. For example, in European models for soil and groundwater related exposure, a total of 19 different pathways is used (Carlon and Swartjes 2007). Some important exposure pathways have been described in previous chapters, i.e., exposure through soil and dust ingestion (see Chapter 6 by Bierkens et al., this book) and exposure through vapour intrusion indoors (see Chapter 10 by McAlary et al., this book). In this chapter, six other important pathways have been described in detail, these are exposure through:

- consumption of vegetables (Section 11.2);
- consumption of animal products (Section 11.3);
- consumption of domestic water and inhalation of volatilised domestic water (Section 11.4);
- inhalation of vapours outdoor (Section 11.5);
- inhalation of dust particles, indoors and outdoors (Section 11.6);
- dermal uptake, via soil material, outdoors and indoors (Section 11.7).

Note that these exposure pathways follow different exposure routes to enter the human body, i.e., oral, inhalation and dermal routes, respectively.

Some of the soil and groundwater related pathways that were mentioned in Carlon and Swartjes (2007) were not considered in this chapter, because they do not play a significant role at the majority of contaminated sites, for example exposure through inhalation of volatilised irrigation water.

The ingestion of soil attached to vegetables, also mentioned in Carlon and Swartjes (2007), is included as part of the pathway “exposure through consumption of vegetables”.

11.1.2 Calculating Exposure

Human exposure through all oral and inhalative exposure pathways described in this chapter follow a similar pattern. This pattern includes the following three steps:

- the transfer of contaminants from one of the mobile phases of the soil (pore water or soil gas) into a so-called contact medium;
- the intake of that contact medium by human beings;
- the uptake of part of the contaminants from the contact medium into the blood stream and target organs and the corresponding excretion of the remaining part of the contaminants.

The dermal contact pathway differs slightly in that intake of the contaminant from the contact medium into the blood stream takes place directly across the skin. The procedures used to calculate exposure and the derivation of input parameters differ between the different exposure pathways and have been described in detail in the Sections 11.2–11.7. In these Sections, general information and mathematical equations are given. Since exposure models can differ substantially with regard to model algorithms and input parameters (see Swartjes (2007) who compared seven European exposure models), examples of mathematical equations are given that generally relate to specific models.

11.2 Exposure Through Consumption of Vegetables

11.2.1 Significance

Human exposure through consumption of vegetables is potentially significant, since contaminants from soil can be taken up and subsequently bioaccumulate in vegetables which are consumed by humans (see Fig. 11.1).

The relevance of the exposure pathway "exposure through vegetable consumption" depends on the land use. Obviously, this pathway is important for the land uses where vegetable consumption is possible, i.e., residential sites, vegetable gardens, and agricultural sites. Vegetables are often grown on contaminated sites, e.g., in city gardens in former industrial or agricultural areas, or in vegetable gardens along railroad tracks. For these reasons, this pathway is included in almost all European exposure models (Carlon and Swartjes 2007), for example in the Dutch

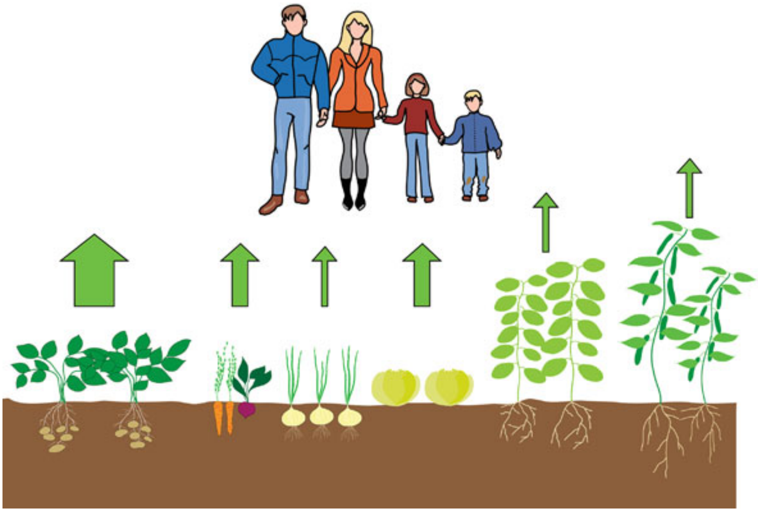


Fig. 11.1 Schematic representation of uptake of contaminants in vegetables and exposure of humans through consumption of these vegetables

CSOIL exposure model (Brand et al. 2007; Van den Berg 1991/1994/1995), the UK CLEA exposure model (DEFRA and EA 2002), the French INERIS' exposure model (INERIS 2003, 2009), the Swedish model for the derivation of Soil Quality Standards (Naturvårdsverket 2009), and the Californian CalTOX exposure model (CalTOX, University of California 1993, California Department of Toxic Substances Control 2009). Moreover, calculated exposure through vegetable consumption is included in almost all human health-based Soil Quality Standards, worldwide. For the derivation of the human health-based Soil Quality Standards in the Netherlands, for example, exposure due to vegetable consumption is important for all metals (Lijzen et al. 2001): the contribution of exposure due to vegetable consumption to total exposure is more than 90% for cadmium and cobalt and between 80% and 90% of the total exposure for copper, mercury, molybdenum and zinc. Exposure through vegetable consumption also contributes a large fraction of the total exposure for several organic contaminants (Lijzen et al. 2001): over 90% for several aromatic contaminants (phenols, catechol, resorcinol and hydroquinone), for some chlorinated contaminants (hexachlorobenzene, pentachlorophenol, several PCBs, chlorophenols) and for several pesticides (DDT, DDE, aldrin, carbofuran, propoxur, atrazine, di(2-ethylhexyl)phthalate). For many dioxins this contribution is close to 80%.

Although the same procedures for Risk Assessment generally apply to crops grown on agricultural sites, this section does not primarily relate to Risk Assessment for this land use. The reason for this is that agricultural produce usually is distributed widely, and vegetables from a contaminated site will be distributed together with vegetables from other sites. Therefore it is unlikely that one person will consume a large fraction of their vegetables from a single contaminated agricultural source. Moreover, contamination of metals, pesticides and nutrients at agricultural sites is often a matter of controlling the addition of contaminants to the site (rate of fertilizer, pesticide and nutrient application, respectively), rather than a question of remediation of contaminated land.

The theory exposed in this chapter, however, is fully applicable to assess the human health risks in the framework of Food Safety.

11.2.2 Conceptual Model

11.2.2.1 Principles

The calculation of exposure through the consumption of vegetables is performed in two stages:

- the calculation of contaminant concentrations in the edible parts of vegetables;
- the calculation of human exposure through consumption of contaminated vegetables.

The calculation of contaminant concentrations in the edible parts of vegetables, the *representative concentration* in vegetables, is the combined result from uptake, transport, accumulation and possibly degradation of contaminants. It is important to realise that the representative concentration in vegetables relates to contaminants in vegetables that originate from soil at the time of harvesting and after food preparation. The uptake routes can be via the roots, via the leaves after rain splash or the dry deposition of soil particles on plant surfaces, and via air. Generally, root uptake is the most important route, especially for contaminants that are mobile within the soil. For contaminants that are immobile in the soil, uptake via the leaves can contribute significantly to the contaminant concentration in leafy vegetables. It is possible to include the relevant routes of uptake and the impact of food preparation in the calculation of exposure. When contaminant concentrations in vegetables are measured, it is very difficult to exclude contaminants taken up via external routes that are not related to the contaminated site. Uptake via the leaves, after atmospheric deposition is usually the most important source of contaminants originating from outside the contaminated site. However, measuring makes it easier to deal with the impact of food preparation (mainly washing).

Calculation of the representative concentration in vegetables is described in detail in other parts of this book (see Chapter 8 by McLaughlin et al., for metals; Chapter 9 by Trapp and Legind, for organic contaminants, this book). Therefore, this section will focus on the calculation of human exposure through consumption of contaminated vegetables only.

11.2.2.2 Differences Between Vegetable Types

Obviously, exposure through consumption of contaminated vegetables depends strongly on the amount of vegetables consumed. The uptake and accumulation of contaminants differs greatly between vegetables. There are plants that can accumulate large amounts of metals in above-ground plants and roots and those that are not effective in the uptake of metals. Generally, fast growing leafy vegetables, like spinach, endive and broccoli, show high metal uptake and accumulation rates. Cadmium, however, shows a relatively high accumulation for all vegetables. In addition, different genotypes of the same crop can exhibit substantial differences in uptake and accumulation rates. Wu et al. (2004), for example, demonstrated differences in cadmium uptake among different cotton genotypes. Specific plant types that readily take up contaminants, like *Brassicaceae* (cabbage), *Papilonaceae* (pods), *Poaceae* (grasses), are even used for phytoremediation (Gawronski 2000). The latter group also includes cereals, but they generally do not hyperaccumulate. Some plants are extremely tolerant for soil metals. These plants grow in areas where natural ores of heavy metals occur in the upper layers of the soil. A well-known example of a heavy metal adapted plant, although not an edible plant, is the zinc violet.

The uptake of organic contaminants is related to the root lipid content, which varies between crops. Concentrations of organic contaminants in vegetables should therefore be normalised to the lipid content of the vegetables (Simonich and Hites 1995). As with metals, the uptake of organic contaminants can also vary between different varieties within certain plant types. As a consequence, there is a large variation in the uptake and accumulation of organic contaminants between vegetables.

11.2.2.3 Representative Concentration in Vegetables

To be able to determine the representative concentration in vegetables, the most relevant vegetables must be selected. For site-specific Risk Assessment it is possible to focus on the vegetables that are actually growing on the specific site at the moment of the assessment. However, this assumption may not be appropriate or practical, for the following reasons:

- These vegetables are not always appropriate for the long term representative concentration, because the type of vegetables might change every few years.
- On many contaminated sites no (representative) vegetables are present at the time when the assessment is made.

In addition, Risk Assessments are often designed to take into account *the possibility* of growing vegetables without experiencing unacceptable adverse effects on human health. This requirement implies that, independent of the vegetables currently growing on the site, attention should be focused on a representative consumption pattern. For the Netherlands, for example, relevant vegetables are potato, carrot, beet, radish, onion, tomato, cucumber, cauliflower, cabbage, lettuce, spinach, endive, french bean, string bean, nave beans, kidney beans and rhubarb. Other edible vegetables in the Netherlands are asparagus, leek, celery, brussel sprouts, eggplant, okra, green pepper, pod, pea, marrow, lentil, courgette, maize, corn and broccoli. Obviously, relevant vegetables significantly vary in different regions in the world.

Note that the calculation of the representative concentration in vegetables offers more flexibility in the choice of vegetables than field measurements. In the latter case, there is no other option than focusing on the type of plant (vegetable or non-vegetable) that is available, unless vegetables are planted and grown specially. When there are no vegetables at all, representative vegetable concentrations can be calculated. If measurements are necessary, for example to take some special site-specific factors into account, and there are no vegetables present, measurements of non-edible plants can be made, assuming that uptake in non-edible plants is related to that in vegetables, or again, vegetables may be grown specially for measurement.

11.2.3 Mathematical Equations

11.2.3.1 Principles

As an example of quantifying exposure through vegetable consumption, the procedure used in the Netherlands is described. This procedure could, however, be used universally. In a general form, the exposure through vegetable consumption is calculated as follows:

$$\mathrm{Exposure}_{\mathrm{vegetables}} = \frac{\sum Q_{\mathrm{vegetable}i} \times C_{\mathrm{vegetable}i} \times f_{\mathrm{home\text{-}grown}} \times f_{\mathrm{bioavailability}}}{W} \tag{11.1}$$

in which

$\text{Exposure}_{\text{vegetables}}$ = exposure due to vegetable consumption [mg/$\text{kg}_{\text{body weight}}$/d]
$Q_{\text{vegetable } i}$ = consumption rate of vegetable i [kg_{dw}/d]
$C_{\text{vegetable } i}$ = contaminant concentration in vegetable i [mg/kg_{dw}]
$\text{f}_{\text{home-grown}}$ = fraction of vegetables that is home-grown [–]
$f_{\text{bioavailability}}$ = correction for relative bioavailability in the human body [–]
W = body weight [$\text{kg}_{\text{body weight}}$]

Ideally, calculation of exposure through vegetable consumption should relate directly to vegetables that are generally grown in home gardens and vegetable gardens. However, statistics on this type of vegetable production are lacking in most countries. For this reason, exposure through vegetable consumption is generally related to the *overall consumption pattern*, i.e., the vegetables that humans consume, independent of their origin (grocery, supermarket or home-grown). In the Netherlands, for example, a consumption-weighed vegetable package has been constructed for Risk Assessment purposes (Swartjes et al. 2007). This package is based on the average consumption rates of 32 vegetable types for all age groups and both sexes. Subsequently, the overall vegetable consumption is considered to be proportional to the contribution of each separate vegetable to the total vegetable consumption rate for adults and children.

11.2.3.2 Metals

For metals, a *vegetable-consumption-rate-weighed BCF* (BioConcentration Factor), $BCF_{\text{veg-cr-weighed}}$, is calculated. This is the relation between the metal concentration in vegetables and the metal concentration in soil. The basis for these vegetable-consumption-rate-weighed BCF are plant-soil relations, i.e., equations which describe the relation between contaminants in specific vegetables as a function of (total) soil concentration and soil properties (for metals; see Chapter 8 by McLaughlin et al., this book). In the Netherlands, a geometric mean of BCFs from the literature is used, which is corrected with a so-called *soil specific correction factor*, in case no significant plant-soil relations are available for a specific vegetable (Swartjes et al. 2007). This *soil specific correction factor* is a function of the organic matter and clay contents of the soil and accounts in a practical way for bioavailability. The equation for the vegetable-consumption-rate-weighed BCF (BioConcentration Factor), $BCF_{\text{veg-cr-weighed}}$, is as follows:

$$BCF_{\text{veg - cr - weighed}} = \Sigma\left(w_i \times BCF_{\text{pl - soil relation}}\right) + \Sigma\left(w_j \times BCF_{\text{geo mean}}\right) / STcf_{BCF} \tag{11.2}$$

in which

w_i = vegetable-consumption-rate-weighing factor, for vegetables for which a plant-soil relation is available [–]

$BCF_{\text{pl-soil relation}}$ = BCF based on plant-soil relation [(mg/kg$_{\text{dw-plant}}$)/(mg/kg$_{\text{dw-soil}}$)]
w_j = vegetable-consumption-rate-weighing factor, for vegetables for which a geometric mean BCF is used [–]
$BCF_{\text{geo mean}}$ = BCF based on geometric mean [(mg/kg$_{\text{dw-plant}}$)/(mg/kg$_{\text{dw-soil}}$)]
$STcf_{BCF}$ = soil specific correction factor [–]

The following nine *vegetable groups* have been distinguished in the Netherlands: Potatoes, Roots and tubers, Bulbous vegetables, Fruiting vegetables, Cabbages, Leafy vegetables, Legumes, Beans, Stem and stalk vegetables. The number of vegetables for which plant-soil relations are available varies widely between the different vegetable groups. Therefore, if plant-soil relations were to be used for each individual vegetable, some vegetable groups with a large number of different types of vegetable would be over-represented in the overall *vegetable-consumption-rate-weighed* BCF. In order to prevent too much weight being given to such groups, the overall consumption-rate-weighed BCF is based on the plant-soil relations, or geometric means, for the *vegetable groups* instead of for the individual vegetables. To this purpose, the BCFs for each vegetable group ($BCF_{\text{vegetable group}}$) are calculated from the BCFs for each separate vegetable within that group and the contribution of that vegetable to total consumption of all vegetables in that group, calculated according to Eq. (11.2). These BCFs for each vegetable can either result from a plant-soil relation or from a geometric mean of measured BCFs with soil type correction. Subsequently, the *vegetable-group-consumption-rate-weighed*, $BCF_{\text{veg-gr-cr-weighed}}$, is calculated as follows:

$$BCF_{\text{veg-gr-cr-weighed}} = \Sigma\,(u_i \times BCF_{\text{vegetable group}}) \tag{11.3}$$

in which

u_i = vegetable-group-consumption-rate-weighing factor [–]
$BCF_{\text{vegetable group}}$ = vegetable-consumption-rate-weighed average BCF for a specific vegetable group [(mg/kg$_{\text{dw-plant}}$)/(mg/kg$_{\text{dw-soil}}$)]

11.2.4 Input Parameters

11.2.4.1 Consumption of Vegetables

Obviously, the type of vegetables that are home-grown strongly depends on the climate and local soil conditions. But even in areas with the same climatic conditions, even within relatively small countries, the home-grown vegetable pattern can vary widely, mainly because of differences in soil properties and region-specific conditions. As a consequence, for every Risk Assessment the national, regional, or local vegetable consumption pattern, i.e., consumption rate of all relevant vegetables, has to be determined.

In most countries, statistics are available of the total vegetable consumption rate and the consumption pattern, although these statistics often are outdated. As mentioned in Section 11.2.2.3, it does not always make sense to relate a vegetable(-group)-consumption-rate-weighed BCF to the vegetables that are actually grown at the site. The reason for this is that this vegetable pattern does not represent a sustainable situation, i.e., the type of vegetables might change the next several years. In most cases it is more useful to consider the risks from the consumption of vegetables produced over a longer time span, where consumption is represented by an average vegetable package for the country or region. An exception could be made when there is a tradition of producing a specific vegetable on the contaminated site, because of either local conditions or regional traditions. In this case, a site-specific or region-specific food package could be assessed, or the general country-specific vegetable consumption statistics could be adapted to the regional conditions. In both cases, the result is a specific BCF for the site or region, as the basis of a "regional" vegetable group-consumption-rate-weighing.

In a screening Risk Assessment, it is possible to focus on vegetables that have a high uptake affinity for the contaminant. When human health risk can be excluded, the Risk Assessment is then concluded. When risks cannot be excluded, more realistic vegetable consumption patterns have to be considered in subsequent stages of the Risk Assessment.

11.2.4.2 Fraction of Vegetables that is Home-Grown

In the majority of countries, the fraction of vegetables that is home-grown, $f_{\text{home-grown}}$, is not specified for each vegetable separately. In most countries, no decent statistics on this fraction exist. Besides, this fraction will differ between regions and between cities and rural areas. An exception are the detailed statistics on the fraction of vegetables that is home-grown for France (INSEE 1991). These fractions are given on the scale of six different regions. Generally, the specification of the fraction of vegetables that is home-grown is partly a judgement, or a policy decision. In the Netherlands, as an example, the following default data are used for this fraction for groups of vegetables, depending on land use:

- "Residential site, with garden": 10% for all vegetables;
- "Vegetable garden": 50% for potatoes and 100% for other vegetables.

The differentiation of the fraction of vegetables that is home-grown between potatoes and other vegetables for the land use "Vegetable garden" is made because it is not realistic to use a value of 100% for potatoes. Because the consumption rate of potatoes is rather high, an exceptionally large vegetable garden would be needed to provide all the potatoes required. These values reflect the policy decision made in the Netherlands that: "the soil quality must offer the possibility of consuming at least a specified percentage of home-grown vegetables". Since the exposure to contaminants is sensitive to the fraction of vegetables that is home-grown, further investigation of the fraction of vegetables that is home-grown has been recommended for the future (Swartjes et al. 2007).

11.2.4.3 Correction for Relative Bioavailability in the Human Body

The relative oral bioavailability of contaminants in the human body (see Chapter 7 by Cave et al., this book, for a detailed explanation) may be an important factor for this pathway, though it is not as important as it is for exposure through the direct ingestion of soil. Intawongse and Dean (2006) studied the oral bioavailability of cadmium, copper, manganese and zinc ingested in the leaves of lettuce and spinach and the roots of radish and carrot and found that the bioaccessibility of metals in plants varies between metals and between different plant types. However, there is only limited information on the differences between intake and uptake of contaminants via a vegetable matrix. Therefore, the correction for bioavailability in the human body, $f_{\text{bioavailability}}$, is currently 1.0 in practically all existing exposure models.

11.2.5 Site-Specific Risk Assessment of Exposure Though Vegetables Consumption

As with other pathways, a tiered approach can be used to assess the site-specific human health risks through vegetable consumption from contaminated sites in a scientifically-based and efficient way. Successively, in each tier the degree of conservatism decreases, while site-specificism increases. As a consequence, complexity and hence effort and finances needed also increase in each tier. In a specific tier, if unacceptable human health risks cannot be excluded, the Risk Assessment should proceed to the following tier. The underlying principle is: simple when possible and complex when necessary.

An example of this kind of tiered approach is given in Swartjes et al. (2007). According to this approach, site-specific Risk Assessment is usually carried out in higher tiers. Site-specific Risk Assessment can take the form of site-specific calculations, or include measurements of metal concentrations in vegetables grown on the site. Tier 0 concerns a preliminary qualitative evaluation of the *possibilities* for experiencing adverse human health effects due to vegetable consumption. In Tier 1 the actual total soil concentrations (average values or a specified percentile of the measured values) are compared with *Critical soil concentrations*. These Critical soil concentrations have been derived on the basis of a conservative exposure scenario. Tier 2 offers the possibility for *a detailed assessment of the site-specific risks on the basis of calculation*. The site-specific calculation of the contaminant concentration in vegetables differs for metals, other inorganic contaminants and organic contaminants. For metals, Freundlich-type plant-soil relations (dependent of the total soil concentration and the major soil properties) (see Chapter 8 by McLaughlin et al., this book) and geometric means of the BioConcentrationFactors (corrected for organic matter and clay contents) are combined. The accumulation of other inorganic contaminants is based on passive uptake. The calculation of the concentration of organic contaminants in vegetables is based on an adapted Trapp and Matthies model (see Chapter 9 by Trapp and Legind, this book). In this model the partitioning of contaminants between pore water and roots and the subsequent translocation to the upper plant parts is calculated, in order to estimate the contaminant concentration in the above-ground plant parts.

Finally, in Tier 3, a *standardised measurement protocol* has been developed. This protocol allows for sampling of a significant number of representative vegetables in the field, for which the edible parts of the plants are treated in the laboratory in analogy with standard kitchen preparation. Subsequently, the measured concentrations can be used in an exposure calculation or, when appropriate, compared to acceptable concentrations in vegetables.

11.2.6 Further Considerations

To assess the risk due to vegetable consumption it is essential to include exposure due to soil ingestion (see Chapter 7 by Bierkens et al., this book) in most cases, because hand-mouth contact is relatively intensive during gardening, also for adults.

People with "Vegetable gardens" often consume more vegetables than the general population. In Swartjes et al. (2007) a 10% higher consumption rate of potatoes was derived for this group of people, in the Netherlands. For most vegetables the consumption rate for people with gardens is approximately 70% higher for adults and schoolchildren and 20% higher for babies and pre-school children.

In many developed countries, a relatively high percentage of vegetables gardens are used by immigrants. In these cases, the quality of a Risk Assessment is improved when the vegetable package relates to the specific vegetables that these immigrants grow.

11.2.7 Reliability and Limitations

In Swartjes (2007) it was demonstrated that the variation in calculated exposure through vegetable consumption between seven European models was substantial, i.e., substantially higher than the variation in calculated exposure through soil ingestion, but much lower than the variation in calculated exposure through indoor air inhalation. It is generally recognised, however, that the most uncertain aspects of this calculation is in the assessment of the representative concentration in vegetables (see Chapter 8 by McLaughlin et al., for metals; Chapter 9 by Trapp and Legind, for organic contaminants, this book).

11.3 Exposure Through Consumption of Animal Products

Human exposure through consumption of animal products from contaminated sites is potentially significant, as some contaminants (for example, dioxins, PCB) can bioaccumulate in the food chain (see Fig. 11.2 for the transfer of contaminants into a cow as an example).

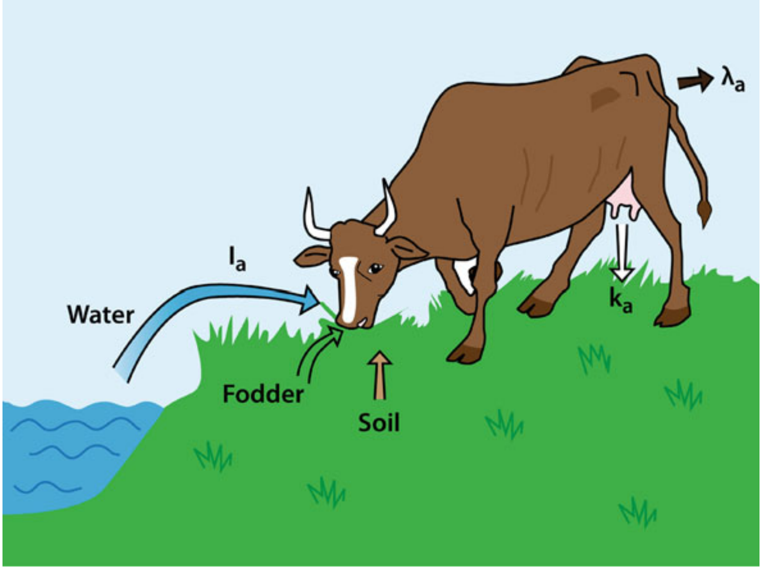


Fig. 11.2 Schematic representation of uptake of contaminants in a cow, as an example of the transfer to an animal product

Depending on the land use and the type of contaminant it may be necessary to take contamination of meat, dairy products and eggs, into account, in order to assess human exposure through food consumption. It may also be necessary to consider the contamination of animal fodder.

This section deals with contamination and consumption of products from terrestrial animals. It does not deal with contamination and consumption of fish or other seafood.

11.3.1 Conceptual Model

Exposure modelling is normally carried out in two stages:

- prediction of contaminant concentrations in the various animal tissues;
- calculation of human exposure through consumption of animal products.

11.3.1.1 Prediction of Contaminant Concentrations in Animal Tissues

Modelling of concentrations in animal products requires the assessment of:

- animal intake (i.e., intake *by* the animal);
- contaminant bioaccumulation in animal tissues.

In principle, animals can be exposed to contaminants through the same pathways as humans. However, the pathways considered for calculating animal exposure are based mainly on ingestion of contaminants: intake of drinking water, direct intake of soil, intake of foodstuffs. Generally, inhalation is neglected. This is because the contaminants for which assessment of risks from consumption of animal products is relevant are persistent and bio-accumulating contaminants, which generally have low volatility.

The bioaccumulation of a contaminant in animal tissues depends on the rates of absorption, degradation and elimination within the animals. For the organic contaminants, the potential for bioaccumulation increases with the hydrophobicity of the contaminant. Lipophilic contaminants tend to accumulate in fat tissue. The uptake of contaminant by animals causes an increase of the quantity of contaminant in the animal tissues up to a steady-state level, at which the contaminant uptake is balanced by degradation and elimination. The lower the degradation and elimination rate, the longer it takes to reach the steady-state and the higher the steady-state concentration will be.

In most Human Health Risk Assessment studies, the contaminant concentrations in animal tissues are considered to be at steady-state. A motive for this is that generally the aim is to assess chronic exposure and the intake by animals is assumed to be constant. Under these conditions the steady state concentration does not underestimate the concentration in animal tissues. In this case, concentrations in animal products (in general, meat, milk, eggs) are calculated by multiplying the daily intake or the animal food concentration by a constant factor called biotransfer or bioconcentration coefficient, respectively. Regression models based on the lipophilicity of contaminant have also been developed to estimate the values of biotransfer coefficient (see Section 11.3.3.2).

However, calculations can also be performed using a dynamic approach. Such an approach is useful under the following conditions:

- to assess concentrations in tissues before the steady-state is reached;
- to estimate the time required for the concentration to revert to a previous level after a period of high exposure or
- in the case of contaminant intake changing with time.

In the simple approach described below, animals are represented by a single compartment and the animal tissue concentration is calculated by a mass-balance between input (uptake) and output (excretion and degradation). For lipophilic contaminants, the production of eggs (for poultry) and milk (for dairy cows) are assumed to be the main elimination pathways.

11.3.1.2 Calculation of Human Exposure

The calculation of human exposure considers the average consumption of animal products from the site over the relevant exposure period. Exposure is calculated by multiplying the concentration of the contaminant in each product by the daily ingestion rate of that product (daily intake in mg contaminant per day), which is then divided by the body weight (daily intake in mg/$\text{kg}_{\text{body weight}}$ and day).

11.3.2 Mathematical Equations

11.3.2.1 Calculation of Animal Intake

Animal intake through food, water and soil ingestion are summed up, as follows:

$$I_a = \sum_j Q_{a,j} \times C_j + Q_{a,s} \times C_s \times B_s + Q_{a,\mathrm{water}} \times C_{\mathrm{water}} \tag{11.4}$$

in which

I_a = daily contaminant intake of the animal [mg/d]
$Q_{a,j}$ = quantity of plant j grown on the contaminated site and ingested by the animal [kg/d]
C_j = concentration of contaminant in plant j [mg/$\mathrm{kg_{dw}}$]
$Q_{a,s}$ = quantity of soil ingested by the animal [kg/d]
C_s = average concentration of contaminant in soil [mg/$\mathrm{kg_{dw}}$]
B_s = relative oral bioavailability [-] in the body of the animal of the contaminant in soil. B_s is the absolute oral bioavailability in the body of the animal of the soil contaminant divided by the bioavailability of the contaminant in food or water. As for humans, B_s is used, because soil contaminants are assumed to be less bioavailable than the same contaminants ingested with food or water.
$Q_{a,\mathrm{water}}$ = quantity of water ingested by the animal [l/d]
C_{water} = concentration of contaminant in water [mg/L]

To follow the variations of the contaminant concentrations in animal products for an annual cycle, animal intake needs to be defined as a function of time. However, within the framework of an evaluation of chronic risks (from one to several years of exposure), the average intake by animals over the exposure duration is often sufficient.

The relevant fodder items will depend on the specific situation at the contaminated site and the conditions of breeding. To assess the accumulated concentration in the animal tissue that originates from the contaminated site, only the fraction of the total dietary intake of the animals (plants) which is grown on the contaminated site and the water impacted by site contamination needs to be considered in the animal diet. In order to assess the total concentration in animal tissue, intake from background should be added.

11.3.2.2 Calculation of the Concentration of Contaminant in Animal Products

Steady-State Approaches

The most common approach consists in multiplying the animal intake by a biotransfer factor (the ratio of animal tissue concentration and the daily intake by the animal):

$$C_{a,i} = I_a \times B_{T\,a,i} \tag{11.5}$$

in which

$C_{a,i}$ = concentration of contaminant in animal tissue i [mg/kg$_{\mathrm{dw}}$]
I_a = average contaminant daily intake for the animal [mg/d]
$B_{Ta,i}$ = contaminant biotransfer factor for animal tissue i (meat, milk, egg) [d/kg]

Depending on the available data, the concentration in animal tissue can also be calculated with a bioconcentration factor (the ratio of animal tissue concentration and the diet concentration):

$$C_{a,i} = \frac{I_a}{\sum Q_{a,j}} \times BCF_{a,i} = \overline{C}_j \times BCF_{a,i} \tag{11.6}$$

in which

$BCF_{a,i}$ = bioconcentration factor for tissue i of animal a [–]
$Q_{a,j}$ = daily quantity of media j ingested by animal a [kg/d]
$\overline{C}_j$ = animal-consumption-rate-weighed media concentration [mg/kg]

$B_{Ta,i}$ and $BCF_{a,i}$ are linked by the following expression:

$$\mathrm{B}_{Ta,i} = \frac{BCF_{a,i}}{\sum_j Q_{a,j}} \tag{11.7}$$

Dynamic Approaches

Meat Concentration

It is assumed that:

- animals can be represented by a unique compartment in which contaminant has a uniform concentration,
- contaminant is eliminated and metabolised according to a first-order reaction rate.

Under these conditions the development of the contaminant concentration in the animal can be described as follows:

$$\frac{dm_{a,1}}{dt} = f_{abs\,a} \times I_a - k_a m_{a,1} - \lambda_a m_{a,1} \tag{11.8}$$

in which

$m_{a,1}$ = mass of contaminant in the animal [mg]
$f_{abs,\,a}$ = absorption fraction of contaminant by the animal [–]
I_a = contaminant daily intake of the animal [mg/d]

k_a = contaminant transfer rate or elimination rate to milk or eggs [d^{-1}]
λ_a = rate of contaminant degradation plus elimination by pathways other than excretion of milk or eggs [d^{-1}]

If I_a is constant with time, the mass of contaminant in animal is as follows:

$$m_{a,1}(t) = \frac{f_{abs,a} \times I_a}{k_a + \lambda_a} \times \left(1 - e^{-(k_a+\lambda_a)\times t}\right) + m_{a,1}(0) \times e^{-(k_a+\lambda_a)\times t} \tag{11.9}$$

and

$$C_{a,1}(t) = \frac{f_{abs,a} \times I_a}{(k_a + \lambda_a) \times M_{a,1}(t)} \times \left(1 - e^{-(k_a+\lambda_a)\times t}\right) + \frac{m_{a,1}(0)}{M_{a,1}(t)} \times e^{-(k_a+\lambda_a)\times t} \tag{11.10}$$

in which

$C_{a,1}(t)$ = contaminant concentration in animal tissues at time t [mg/kg]
$M_{a,1}(t)$ = weight of the animal tissues at time t [kg]

If $m_{a,1}(0) = 0$ then,

$$C_{a,1}(t) = \frac{f_{abs,a} \times I_a}{(k_a + \lambda_a) \times M_{a,1}(t)} \times \left(1 - e^{-(k_a+\lambda_a)\times t}\right) \tag{11.11}$$

When the contaminant is mainly accumulated in fat tissue, the fat concentration can be calculated from the previous equations ($C_{a,1}(t)$ = contaminant concentration in fat tissues and $M_{a,1}(t)$ = weight of fat tissue in the animal at time t) and the meat concentration can be derived from $C_{a,1}$ by multiplying it by the fraction of fat in meat.

If I_a varies with time, $m_{a,1}$ has to be integrated on time intervals ΔT where I_a may be regarded as constant:

$$m_{a,1}(t) = \frac{f_{abs,a} \times I_a}{(k_a + \lambda_a)} \times \left(1 - e^{-(k_a+\lambda_a)\times \Delta T}\right) + m_{a,1}(t - \Delta T) \times e^{-(k_a+\lambda_a)\times \Delta T} \tag{11.12}$$

and

$$C_{a,1}(t) = \frac{m_{a,1}(t)}{M_{a,1}(t)} \tag{11.13}$$

When for very lipophilic contaminants there is neither metabolism, nor elimination via milk or eggs (e.g., in case of beefs or chickens), concentrations can be assessed as follows, as elimination via urine is assumed to be low:

$$C_{a,1}(T) = \frac{\sum_{t_i=0}^{T} [I_a(t_i) \times (t_i - t_{i-1})] \times f_{abs,a}}{M_{a,1}(T)} \tag{11.14}$$

Or, if I_a is constant

$$C_{a,1}(T) = \frac{I_a \times T \times f_{abs,\,a}}{M_{a,1}(T)} \tag{11.15}$$

in which

$T =$ exposure time [d]

Milk and Egg Concentrations

Contaminant mass in milk or eggs is calculated from the following differential equation:

$$\frac{dm_{a,2}}{dt} = k_a \times m_{a,1} \tag{11.16}$$

in which

$m_{a,2}$= total mass of contaminant excreted by the animal (i.e., mass of contaminant in milk produced by cows or eggs produced by hens) [mg]

If I_a is constant with time, the calculation of the excreted contaminant mass is as follows:

$$\frac{dm_{a,2}}{dt} = \frac{k_a}{k_a + \lambda_a} \times f_{abs,\,a} \times I_a \times \left(1 - e^{-(ka+\lambda_a)\times t}\right) + k_a \times m_{a,1}(0) \times e^{-(ka+\lambda_a)\times t} \tag{11.17}$$

The solution of this equation can be written as follows:

$$\begin{aligned} m_{a,2}(t) = {} & \tfrac{k_a}{k_a+\lambda_a} \times f_{abs,\,a} \times I_a \times t + \tfrac{k_a}{(ka+\lambda_a)^2} \times f_{abs,\,a} \times I_a \times e^{-(ka+\lambda_a)\times t} \\ & -\tfrac{k_a}{k_a+\lambda_a} \times m_{a,1}(0) \times e^{-(ka+\lambda_a)\times t} + \text{Constant} \end{aligned} \tag{11.18}$$

As $m_{a,2}(0) = 0, m_{a,2}(t)$, can be rewritten as follows:

$$\begin{aligned} m_{a,2}(t) = {} & \tfrac{k_a}{k_a+\lambda_a} \times f_{abs,s} \times I_a \times t + \tfrac{k_a}{(k_a+\lambda_a)} \left[m_{a,1}(0) \text{ - } \tfrac{f_{abs,s} \times I_a}{(k_a+\lambda_a)} \right] \\ & \times \left(1 - e^{-(k_a+\lambda_a)\times t}\right) \end{aligned} \tag{11.19}$$

and

$$C_{a,2}(\Delta T) = \frac{m_{a,2}(t) - m_{a,2}(t - \Delta T)}{M_{a,2}(t) - M_{a,2}(t - \Delta T)} \tag{11.20}$$

in which

$C_{a,2}(\Delta T)$ = average contaminant concentration in the excreted tissue (milk or eggs) by the animal for ΔT [mg/kg]
$M_{a,2}$ (t) = mass of tissue excreted by animal a from t=0 to t [kg]

If ΔT is equal to one day and $M^j_{a,2}$ (daily weight of tissue produced by the animal) is constant,

$$C_{a,2}(t) = \tfrac{k_a}{k_a+\lambda_a} \times \tfrac{f_{abs,s} \times I_a}{M^j_{a,2}} + \tfrac{k_a}{(k_a+\lambda_a)} \times \left[m_{a,1}(0) - \tfrac{f_{abs,s} \times I_a}{(k_a+\lambda_a)}\right] \times \frac{\left[e^{-(k_a+\lambda_a)\times(t-1)} - e^{-(k_a+\lambda_a)\times t}\right]}{M^j_{a,2}} \tag{11.21}$$

If I_a varies with time, $m_{a,2}$ has to be integrated over short time intervals where I_a and $m_{a,1}$ may be regarded as constant:

$$m_{a,2}(t) = m_{a,2}(t - \Delta T) + \Delta T \times \tfrac{k_a}{(k_a+\lambda_a)} \times f_{abs,s} \times I_a \times \left(1 - e^{-(k_a+\lambda_a)\times\Delta T}\right) + \Delta T \times k_a \times m_{a,1}(t - \Delta T) \times e^{-(k_a+\lambda_a)\times\Delta T} \tag{11.22}$$

If ΔT is equal to one day,

$$C_{a,2}(t) = \frac{m_{a,2}(t) - m_{a,2}(t - \Delta T)}{M^j_{a,2}} \tag{11.23}$$

and

$$C_{a,2}(t) = \frac{\tfrac{k_a}{k_a+\lambda_a} \times f_{abs,s} \times I_a \times \left(1 - e^{-(k_a+\lambda_a)}\right) + k_a \times m_{a,1}(t-1) \times e^{-(k_a+\lambda_a)}}{M^j_{a,2}} \tag{11.24}$$

Time Required to Reach Steady-State Concentrations and Definition of B_{Ta}

According to equation (11.11), when growth of animals is negligible and I_a is constant, $C_{a,1}$ is close to the steady-state if:

$$C_{a,1}(t) \approx \frac{f_{abs,s} \times I_a}{(k_a + \lambda_a) \times M_{a,1}} \tag{11.25}$$

or

$$e^{-(k_a+\lambda_a)t} \approx 0 \tag{11.26}$$

With

$$t \geq \frac{3}{k_a + \lambda_a}, e^{-(k_a+\lambda_a)t} < 0,05 \tag{11.27}$$

Thus, if animals have a constant exposure level for a period superior to $3/(k_a+\lambda_a)$ then the steady-state approach based on biotransfer coefficients can be used to estimate meat concentration.

According to equations (11.5) and (11.25), the biotransfer coefficient for animal's tissues is given by:

$$B_{Ta,1} = \frac{f_{abs,\,a}}{(k_a + \lambda_a) \times M_{a,1}} \tag{11.28}$$

If the contaminant is mainly present in fat tissues and $B_{Ta,1}$ is the biotransfer coefficient for fat (in this case $M_{a,1}$ = weight of fat tissues in the animal), the biotransfer coefficient in meat can be calculated by multiplying Eq. (11.28) by the fat fraction in meat.

The biotransfer coefficient for the excreted tissue (milk or eggs) can be deduced from equation (11.16):

$$C_{a,2} \times \frac{dM_{a,2}}{dt} + M_{a,2}\frac{dC_{a,2}}{dt} = k_a \times m_{a,1} \tag{11.29}$$

If $M_{a,2}(t) = M^j_{a,2} \times t$ and $M_{a,1}$ constant then,

$$\frac{dC_{a,2}}{dt} = \frac{k_a \times M_{a,1} \times C_{a,1}}{M^j_{a,2} \times t} - \frac{C_{a,2}}{t} \tag{11.30}$$

and at the steady state:

$$C_{a,2} = \frac{k_a \times M_{a,1} \times C_{a,1}}{M^j_{a,2}} \tag{11.31}$$

Rearranging equations (11.31) with equation (11.25) yields:

$$C_{a,2} = \frac{k_a}{(k_a + \lambda_a)} \times \frac{f_{abs,a} \times I_a}{M^j_{a,2}} \tag{11.32}$$

and

$$B_{Ta,2} = \frac{k_a}{(k_a + \lambda_a)} \times \frac{f_{abs,\,a}}{M^j_{a,2}} \tag{11.33}$$

11.3.2.3 Calculation of Human Exposure

The exposure is calculated as follows:

$$D_f(T) = \frac{1}{T}\sum_{t_i=0}^{T}\sum_{j}\left[Q_j(t_i) \times C_j(t_i) \times f_i(t_i)\right] \times \frac{1}{BW(t_i)} \times (t_i - t_{i-1}) \tag{11.34}$$

in which

$D_f(T)$ = average human exposure to contaminant by ingestion of animal products for the period from $t = 0$ to T [mg/kg$_{\text{body weight}}$/d]
$Q_j(t_i)$ = daily average quantity of product j consumed at time t_i [in kg/d]
$C_j(t_i)$ = concentration of contaminant in product j at time t_i [mg/kg]
$f_j(t_i)$ = fraction of products j from the site at time t_i [–]
$BW(t_i)$ = body weight of the receptor at time t_i [kg]

For acute exposure, T is between 1 to several days, for subchronic exposure T is between several days to several months and for chronic exposure T may range from one year to a lifetime.

If the input parameters of Eq. (11.34) can be assumed constant during T, the average human exposure by ingestion of animal products is as follows:

$$D_f = \frac{\sum_{j} Q_j \times C_j \times f_j}{BW} \tag{11.35}$$

11.3.3 Input Parameters

11.3.3.1 Intake by Animals

Data from different guidelines, reports and scientific papers are gathered in the Tables 11.1, 11.2, and 11.3.

Some references distinguish the diet of the dairy cows from that of the beef cattle, as the former have a higher consumption of food and water, which can lead to greater exposure to contaminants.

Soil ingestion can be due to inadvertent ingestion of soil while eating outdoors or to deposition of soil on forage. That is why soil ingestion is sometimes expressed as a fraction of the consumption of forage by animals. There is little information about soil ingestion rates, therefore the uncertainties in the values are relatively high.

As the absorption rate or biotransfer parameters of contaminants are usually estimated from experiments where animals were exposed via food or water, the relative oral bioavailability (B_S) of contaminants in these experiments is not necessarily the same as the relative oral bioavailability of contaminants on soil particles.

Table 11.1 Total food intake of animals [kg/d]

References	Cattle	Pigs	Lambs	Poultry
Veerkamp and Ten Berge (1994) [kg dry matter/d]	16.5			0.1
IAEA (1994) [kg dry matter/d]	Dairy cows: 16.1 (10–25) Beef cattle : 7.2 (5–10)	2.4 (2–3)	1.1 (0.5–2.0)	Laying hens: 0.1 (0.07–0.15) Chickens: 0.07 (0.05–0.15)
IAEA (2001) [kg dry matter/d]	Dairy cows: 16 Beef cattle: 12			
GCNC (2002) [kg fresh matter/d]	Grass: 60 (8 month/year) Hay: 5 (4 month/year) Corn: 35 (4 month/year)	Corn silage: 2.5	Grass: 8	Corn silage: 0.06
US DOE (2003) [kg fresh matter/d]	Beef cattle: 29–68 Dairy cows: 50–73			0.11–0.4
US EPA (2005) [kg dry matter/d]	Forage (grass/hay): Beef cattle: 8.8 Dairy cows: 13.2 Silage: Beef cattle: 2.5 Dairy cows: 4.1 Grain: Beef cattle: 0.47 Dairy cows: 3.0	Silage: 1.4 Grain: 3.3		Grain: 0.2

Depending on the source, food intake rates are either given on a fresh matter basis or a dry matter basis. The dry to wet-weight ratios are around 20, 90,18 and 90% for grass, hay, corn silage and grains respectively (GCNC 2002; US DOE 2004)

Differences in the relative oral bioavailability, the variable Bs, should be taken into account when assessing the internal exposure through soil ingestion (see Chapter 7 by Cave et al., this book). In the absence of reliable data, B_S is often assumed to be equal to 1, which is a conservative approach.

Table 11.2 Water intake of animals [L/d]

References	Cattle	Pigs	Lambs	Poultry
Veerkamp and Ten Berge (1994)	55			0.2
IAEA (1994)	Dairy cows: 50–100 Beef cattle: 20–60	6–10	3–5	0.1–0.3
IAEA (2001)	Dairy cows: 60 Beef cattle: 40			
GCNC (2002)	18	10	4	0.2
US DOE (2003)*	Beef cattle: 60 Dairy cows: 60–100			0.5

*Data given by US DOE correspond to a climate with high temperature

Table 11.3 Soil intake rates of animals [kg/d]

References	Cattle	Pigs	Lambs	Poultry
Veerkamp and Ten Berge (1994) [kg/d]	0.72	0		2.10^{-3}
IAEA (1994) [%]	Grazing cattle: 6% of dry weight of feed	1.5% of dry weight of feed (if confined feed)	10 % of dry weight of feed	
GCNC (2002) [kg/d]	0,7	0	0.32	10^{-2}
GRNC (2002) [kg/d]	0.1–0.5			
US DOE (2003) [kg/d]	Beef cattle: 0.4–1.0 Dairy cows: 0.8–1.1			10^{-2} to 3.10^{-2}
US EPA (2005) [kg/d]	Beef cattle: 0.5 Dairy cows: 0.4	0.37		2.10^{-2}

11.3.3.2 Parameters for Estimating the Concentration in Animal Tissues

To calculate the concentration of a contaminant in animal tissues it is necessary to know the oral absorption, elimination and degradation rates, or the biotransfer or bioconcentration factors. The choice of values for these parameters is a very important step in the modelling. These parameters may vary with the animal (type of animal, age, gender) and with the contaminant (type of contaminant, speciation, dose). The determination of these parameters, which is done by performing experiments on domestic animals, is relatively difficult, time-consuming and, therefore, expensive. As a consequence, data for these input parameters are rare and uncertain. Values collected from the literature may vary over several orders of magnitude. When empirical data are lacking, regression models are often used to predict values of bioconcentration or biotransfer coefficients. To reduce uncertainties on biotransfer and bioconcentration coefficients, data and regression relationships collected from literature have to be critically reviewed in order to make sure that the experimental protocol used to generate the data was relevant and that the data are of good quality.

Method for Estimating Depuration and Absorption Rates

From the last day of animal exposure, the depuration rate can be estimated from the slope of the curve where the concentration in tissue is plotted against time.

When $I_a = 0$ and $M_{a,1}$ is constant, this yields:

$$\frac{dC_{a,1}}{dt} = -\Lambda \times C_{a,1} \tag{11.36}$$

in which

$\Lambda = k_a + \lambda_a =$ depuration rate [d^{-1}]

$$lnC_{a,1}(t) = -\Lambda \times t - \text{constant} \tag{11.37}$$

The absorption rate is calculated from data measured during the exposure period, as follows:

$$f_{abs,a} = \frac{\frac{dm_{a,1}(t)}{dt} + \Lambda \times m_{a,1}(t)}{I_a(t)} \tag{11.38}$$

or, if $M_{a,1}$ can be assumed to be constant:

$$f_{abs,a} = \frac{\left(\frac{dC_{a,1}(t)}{dt} + \Lambda \times C_{a,1}(t)\right) \times M_{a,1}}{I_a(t)} \tag{11.39}$$

If $\lambda_a \neq 0$, k_a and λ_a can be identified at the steady-state with Eq. (11.32) or by measuring the mean concentration in the excreted tissue from t_1 to t_2, after the end of the exposure period, by using the following equations:

When $I_a\ (t) = 0$, the development of the concentration in excreted tissue is as follows:

$$\frac{dm_{a,2}}{dt} = k_a \times m_{a,1} \tag{11.40}$$

$$m_{a,2}(t) = -\frac{k_a}{k_a + \lambda_a} \times m_{a,1}(T) \times e^{-(k_a+\lambda_a)\times(t-T)} + \text{constant} \tag{11.41}$$

in which

T = last day of exposure [d]

$$m_{a,2}(t_2) - m_{a,2}(t_1) = \frac{k_a}{k_a + \lambda_a} \times m_{a,1}(T) \times \left[e^{-(k_a+\lambda_a)\times(t_1-T)} - e^{-(k_a+\lambda_a)\times(t_2-T)}\right] \tag{11.42}$$

and if $M^j_{a,2}$ is constant:

$$m_{a,2}(t_2) - m_{a,2}(t_1) = M^j_{a,2}\,(t_2 - t_1) \times \overline{C\ {}_{\underset{t_1\to t_2}{a,2}}} \tag{11.43}$$

in which

$\overline{C\ {}_{\underset{t_1\to t_2}{a,2}}}$ average concentration in the excreted tissue from t_1 to t_2

$$k_a = \frac{\lambda_a \times M^j_{a,2}\,(t_2 - t_1) \times \overline{C\ {}_{\underset{t_1\to t_2}{a,2}}}}{m_{a,1}(T) \times \left[e^{-(k_a+\lambda_a)\times(t_1-T)} - e^{-(k_a+\lambda_a)\times(t_2-T)}\right] - M^j_{a,2}\,(t_2 - t_1) \times \overline{C\ {}_{\underset{t_1\to t_2}{a,2}}}} \tag{11.44}$$

Quality Criteria for Designing a Measurement Protocol or Selecting Empirical Data from the Scientific Literature

The following issues should be investigated in order to decide on the suitability of data for calculating the exposure through animal products for the contaminated site under investigation:

- The duration of the experiment. To define a biotransfer or bioconcentration coefficient directly from the concentration in the studied tissue and animal intake, the animal should be exposed to the contaminant for a period that is long enough to enable the concentration in the tissue to reach steady-state. When degradation and elimination rates are very low, steady-state may not be reached before the slaughter of animal. In this case, an apparent B_{Ta} can be calculated as the ratio between the tissue concentration and the daily intake.
- The number and types of animals used in the experiment. Because of inter-individual variability, values based on the same type of animals as those at the contaminated site under investigation and values based on several individuals should be preferred.
- The medium with which the contaminant is administrated to the animals (pure, mixed with food, with soil or with sludge). The oral bioavailability in the body of the animal depends on the medium, and therefore the medium should be representative of the case study, otherwise a relative bioavailability parameter should be introduced to correct the biotransfer or bioconcentration factor.
- The dose of contaminant administrated to the animals. Data should correspond to representative exposure levels, because the animal response may depend on the administrated dose. It means that B_{Ta} or BCF_a are not necessary linear with contaminant intake.
- The presence of potential sources of contamination, other than the dose intentionally applied to animals (including background contamination). If such sources exist, they have to be controlled. A control group of animals is useful to estimate the impact of these sources on concentrations in animal products.
- The measured contaminant concentrations in the administered media must to be greater than to the quantitative limits of the analysis methods used. Using the quantitative or detection limit as the actual concentration in the animal tissue may result in overestimation of B_{Ta} and BCF_a.

Regression Models Available for Predicting Biotransfer Coefficients for Contaminants in Meat and Milk

The most widely used approach for predicting contaminant's biotransfer coefficients in meat and milk are the Travis and Arms (1988) equations. They developed linear regression equation based on octanol-water partition coefficient (*Kow*). For biotransfer coefficient in beef meat this relation is as follows:

$$\log B_{T,1} = -7.6 + \log Kow\, n = 36, r = 0.81 \qquad (11.45)$$

with log *Kow* going from 1.3 to 6.9.

For the biotransfer coefficient in milk the relation is as follows:

$$\log B_{T,2} = -8.1 + \log Kow\, n = 28, r = 0.74 \tag{11.46}$$

with log *Kow* going from 2.8 to 6.9.

Outside these bounds, the model is not considered to be appropriate. In order to calculate B_T in these cases, use of the value calculated with the minimum or the maximum value of Kow for which the model is defined is recommended (European Commission 2008).

The user of Eq. (11.46) should also check that the output rate of contaminant from milk or eggs calculated with this predicted $B_{T,2}$ is less than the intake of contaminant by the animals.

These relationships are used in different multimedia models (CalTOX, University of California 1993, California Department of Toxic Substances Control, 2009; HESP, Veerkamp and Ten Berge 1994; EUSES, European Commission 2008). However, the validity of these relationships has been questioned in the past years, for the following reasons (RTI 2005):

- B_T for contaminants with high *Kow* are overestimated with these relationships (the predicted amount of contaminants in animal products is greater than the amount ingested by the animals),
- they are not suitable for rapidly metabolising contaminants,
- they are based on a mixture of steady-state and non-steady-state data
- they are based on a mixture of data from lactating and non lactating animals

Other relationships have been developed, but these are used less often. An example is the relationships of Dowdy et al. (1996), based on molecular topology (molecular connectivity index) instead of the octanol-water partition coefficient. However, these kind of relations do not appear to be more realistic.

RTI (2005) developed a new relationship based on a methodology which aims to avoid the limits and drawbacks highlighted in the work of Travis and Arms (1988). In their analysis, the authors found no significant difference between contaminant concentrations in milk fat and beef the following equation was derived to predict contaminant concentration in fat:

$$\log B_{T,\text{fat}} = -0.099 \times (\log Kow)^2 + 1.07 \times \log Kow - 3.56,\ r^2 = 0.83 \tag{11.47}$$

$B_{T,\text{meat}}$ and $B_{T,\text{milk}}$ are calculated by multiplying $B_{T,\text{fat}}$ by the fat content of meat and milk, respectively. The authors claimed that this equation is appropriate for organic contaminants with a log *Kow* between –0.67 and 8.2. As with the regression models shown above, for contaminants with log Kow values lying outside this range, $\mathrm{B}_{\mathrm{T,fat}}$ has to be calculated with the minimum or maximum value of Kow for which the model is defined. However, because of the way the relationship was established, it may overestimate biotransfer coefficient for contaminants that are oxidizable or hydrolyzable. If known, a metabolism factor can be applied to improve concentration predictions. However, it should be noted that the metabolised forms of a contaminant are sometimes as toxic as, or even more toxic than the parent compound.

11.3.3.3 Human Consumption of Products

Product Consumption

For site-specific Risk Assessment, region specific or country specific data must be used for the diet, since cultural and climatic differences have a large influence on this parameter. The European Exposure Factors Sourcebook (ExpoFacts 2006) gathers exposure data from 30 European countries. Data regarding diet are defined according to country, region, age, gender and social economic criteria. For some countries, statistical distributions of food consumption rates and patterns are provided.

Depending on the site and the level of detail required in modelling, several categories of data on the human consumption of animal products has to be gathered. Dairy products, eggs and meat are usually taken into account. Because of the differences in exposure conditions between animals, meat from beef, pork and lamb are sometimes distinguished.

For lipophilic contaminants, such as dioxins, which tend to accumulate in fat, it can be more relevant to sum the quantities of products from one category on the basis of their fat content. For these contaminants, dairy products have very different contaminant concentrations when expressed on a fresh weight basis as their fat contents range from 2 or 3% (milk) to more than 80% (butter).Therefore it is preferable to calculate the weight of fat consumed via the different animal products and to multiply it by the contaminant concentration in fat.

Fractions Related to on the Site

To assess human health risk due to a contaminated site, only the fraction of food produced on the site must be taken in account. The fraction of the diet which has a local origin varies substantially between farmer and non-farmer families (see Table 11.4 for data for France, as an example).

11.3.4 Reliability and Limitations

It is generally recognised that the calculated concentrations in animal products are relatively uncertain. The approaches proposed in this Section are simple. They are based on models where animals are represented by only one box. This high simplification assumes a single, uniform concentration in tissues, while in reality the contaminants are heterogeneously distributed throughout the tissues. More sophisticated models exist, like the model of McLachlan (1994), but they require more data that may not be readily available.

Table 11.4 Consumption and subsistence farming of plants and animal products in France (calculated from INSEE 1991)

	Agricultural population		Non agricultural population	
	Consumption [kg/person/year]	Subsistence farming [%]	Consumption [kg/person/year]	Subsistence farming [%]
Vegetables	115.8	67.9	93.0	25.7
Fruit*	47.9	28.0	38.5	12.0
Beef	18.1	32.3	15.1	3.3
Pork	24.15	27.5	19.9	3.5
Poultry	17.5	68.3	13.2	15.3
Eggs$	165.5	58.7	138.0	16.9
Dairy products	121.7	18.7	106.7	2.3

*Excluding tropical fruit
$Consumption expressed in unit per year

Uncertainties and variability related to transfer parameters are large, as described in Section 11.3.3.2. If a regression relation is used instead of empirical data, the uncertainties linked to the definition of the equation have to be added.

In case of a linear regression ($y = ax + b$), like for the relation of Travis and Arms (1988) (where $\log B_T = y$ and $\log Kow = x$), the (1–α) confidence interval of $\log B_T$ for a new contaminant is given by:

$$\left[a \times x^* + b \pm t_\alpha \times \sqrt{\frac{\hat{S}^2 \times (n+1)}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \right] \tag{11.48}$$

in which

x^* =log Kow for a new contaminant
$\bar{x}$ = average of the x_i observed in the dataset used to define the regression relationship
t_α = value of t-distribution with n–2 degrees of freedom, at α level
n = number of (y_i, x_i) in the dataset used to define the regression relationship [–]
$\hat{s}$ = residual standard deviation

There are also uncertainties linked to the assessment of animal exposure, because:

- Some phenomena are not taken into account in models, like ingestion of earthworms by hens which are kept (partly) outdoors.

- Some exposure parameters are poorly known, such as the quantities of soil ingested by animals. Depending on the conditions in which the animals are reared, this parameter can vary substantially (for example, there are differences between soil with a dense cover of grass versus bare ground, and between grazing animals versus confined animals). Exposure data should be adapted to the site under consideration, but reliable reference data are rather scarce, especially for soil ingestion.

Finally, There are uncertainties involved in the calculation of human exposure. They are due to the fact that humans eat different products, which are merged into a few categories, which are more or less representative of the actual consumption and which are assumed to have a single contaminant concentration. Besides, the impact of food preparation on the concentration in animal products is not taken into account. Another factor concerns inter-individual differences between humans. Some individuals consume more animal products concentrating high levels of contaminants (such as offal) than the average population. If such subpopulations are present on the site, their specific exposure should be considered.

11.4 Exposure Via Domestic Water

Groundwater used as domestic water can be contaminated through leaching of contaminants from the soil and subsequent transport to the groundwater. The process of groundwater contamination is very site-specific and is often treated in a simplified manner in Risk Assessment models, e.g., the models used for Soil Quality Standards in USA, Sweden and Flanders. The contaminant concentration in the groundwater is usually calculated based on dilution of the soil pore water in the groundwater, neglecting the temporal evolution of the contaminant plume. See Chapter 19 by Rolle et al., this book. In many countries Soil Quality Standards do not consider groundwater as a protection target for contaminants from the upper soil layers, e.g., the Netherlands and Denmark. Instead separate standards are developed for groundwater quality based on the use of groundwater for domestic purposes and ecological effects in the groundwater.

Domestic water can also be contaminated due to permeation of organic contaminants through plastic water pipes placed in contaminated soil. For example, this pathway is included in the Dutch CSOIL model (Brand et al. 2007; Van den Berg 1991/1994/1995) and the Flemish Vlier-Humaan model (OVAM 2004). Surface water may also be used as a source for domestic water. However, this pathway is rarely considered in Risk Assessment models for contaminated soils, since the groundwater pathway usually gives higher exposure than the surface water pathway, due to less dilution.

This section will only deal with methods used to estimate the exposure due to use of contaminated domestic water, not how the domestic water has been contaminated.

11.4.1 Conceptual Model

Contaminated domestic water can lead to the exposure of humans in several ways:

- consumption of drinking water;
- inhalation of volatilised domestic water;
- dermal contact during showering or bathing.

Normally, consumption of drinking water is the most important of these three pathways. The exposure is directly calculated from the contaminant concentration in the domestic water and the consumption rate of drinking water, either as beverages or in food. Treatment of groundwater may reduce contaminant concentrations before its use as drinking water, but this is normally not taken into consideration in Risk Assessment models.

Volatile contaminants in domestic water will evaporate during tapping or when water is held in open containers, such as water heaters, dishwashers, toilets, bath tubs and washbasins. This exposure pathway has been shown to be of importance for radon present in groundwater. Radon has a very high Henry's Law constant and a relationship has been found between the radon content in domestic water and the radon concentration in indoor air (Gesell and Pichard 1978; Hess et al. 1982). Exposure occurring during showering is often chosen as the critical pathway for volatilization from contaminated domestic water. The relative source strength from showers, considering overall volatilisation efficiency and water consumption, is estimated to be higher for showers than other domestic sources (Howard-Reed et al. 1999). During showering the volatile contaminants can evaporate from the tap water and be inhaled with the water vapour or as fine aerosols. The evaporation rate in a shower is high, since the water droplets have a high surface-volume ratio. Also the elevated temperature of a normal shower will increase the evaporation rate. For contaminants with a dimensionless Henry's Law constant greater than 1, more than 70% of the contaminant in the water is expected to evaporate (Moya et al. 1999). Human exposure is calculated from the concentration of contaminant in the shower stall and the exposure time. Showering may also lead to an increased contaminant concentration in other rooms in the house.

Dermal uptake during showering or bathing has little influence on total human exposure, but is considered in several models, e.g., the Dutch and Flemish models. Dermal exposure is determined by the concentration of the contaminant in the water, the dermal absorption rate, the surface area of the skin that is exposed and the exposure time.

11.4.2 Mathematical Equations

11.4.2.1 Consumption of Drinking Water

The modelling approach for consumption of drinking water is straightforward, where the exposure is given by:

$$E_{dw} = \frac{C_{dw} Q_{dw}}{BW} \tag{11.49}$$

in which

E_{dw} = the exposure due to consumption of drinking water $[\text{mg}\cdot\text{kg}_{\text{body weight}}{}^{-1}\cdot\text{d}^{-1}]$
C_{dw} = the contaminant concentration in the drinking water $[\text{mg}\cdot\text{L}^{-1}]$
Q_{dw} = the average daily consumption of drinking water $[\text{L}\cdot\text{d}^{-1}]$
BW = the body weight [kg]

Since exposure is often calculated for different age groups, age dependent consumption rates and body weights may be needed.

11.4.2.2 Inhalation of Volatilised Domestic Water

In the US EPA Preliminary Clean-up Goals (US EPA 1991) volatilisation from all uses of household water is considered (e.g., showering, laundering, dish washing). An empirical expression is used for the relationship between the concentration of a contaminant in household water and the average concentration of the volatilised contaminant in air (Andelman 1990). This is based primarily on experimental data on the volatilization of radon from domestic water. The equation uses a default "volatilization" constant (K) upper-bound value of 0.5 mg m^{-3} per mg L^{-1}. It is assumed that the volume of water used in a residence for a family of four is 720 L.d^{-1}, the volume of the dwelling is 150 m^3 and the air exchange rate is 0.25 m^3 h^{-1}. Furthermore, it is assumed that the average transfer efficiency weighted by water use is 50% (i.e., half of the concentration of each contaminant in water will be transferred into air by all water uses, the range reported to be from 30% for toilets to 90% for dishwashers).

Special models have been developed for evaluating the evaporation rate of the volatile contaminants during showering by calculating the mass transfer in the boundary layer between the shower water drops and the surrounding air using two-film theory (Foster and Chrostowski 1986; Little 1992). The mass transfer is dependent on the drop size, the fall time of the droplet, the temperature corrected Henry's Law constant and a contaminant specific mass transfer resistance. The mass transfer resistance takes into account the resistance in the water boundary layer of the drop and the air boundary layer surrounding the drop. This type of model is included in the Dutch (CSOIL) and Flemish models (Vlier-Humaan) used for the derivation of Soil Quality Standards.

11.4.2.3 Dermal Contact During Showering

Dermal uptake of contaminants while bathing or showering is modelled using a dermal absorption rate or dermal permeability coefficient. These can either be based on empirical data or derived from empirical relationships using the octanol-water partitioning coefficient (Kow) and the molecular weight of the contaminant. The Dutch CSOIL model uses a single expression for all organic contaminants, while the US Risk Assessment Guidelines for Superfund (US EPA 2004b) uses empirical values for inorganic contaminants and an empirical expression based on Kow and molecular weight for the organic contaminants.

The exposure due to dermal contact further depends on the concentration in the water, the body surface of a child/adult, the fraction of exposed skin during showering or bathing, the exposure time, the relative sorption factor and the body weight.

11.4.3 Input Parameters

11.4.3.1 Consumption of Drinking Water

The intake of drinking water depends on age, climate (intake increases significantly at temperatures over 25°C), physical activity and cultural differences. Water can be consumed directly, in the form of other drinks (e.g., tea, coffee, soup) or indirectly in connection to food preparation. Bottled water can contribute to the total consumption. WHO's drinking water guideline values for potentially toxic contaminants (WHO 1993 and updates, WHO 2006) are based on an intake of 2 L/d for an adult weighing 60 kg, 1 L/d for children with a body weight of 10 kg and 0.75 L/d for bottle-fed babies with a body weight of 5 kg. Children have a greater intake per kg body weight and are more sensitive to many contaminants.

Many models, including the European model for Risk Assessment of chemicals (ECB 2003) and CSOIL (Brand et al. 2007) use the drinking water consumption rates suggested by WHO. A number of national surveys have shown that the average intake is less than the WHO-values, though it varies widely between individuals.

In the IEUBK model for calculating the exposure of children to lead (US EPA 2002b), the US EPA uses age specific drinking water ingestion rates, see Table 11.5.

ECETOC (2001) recommends an average intake of drinking water of 1.1 L/d for adults and 0.5 L/d for children in the age 1–11 years, based on UK studies.

Table 11.5 Age dependent drinking water ingestion rates used in the IEUBK-model

Age [years]	Drinking water ingestion rate [L·d^{-1}]
0–1	0.20
1–2	0.50
2–3	0.52
3–4	0.53
4–5	0.55
5–6	0.58
6–7	0.59

US EPA (2002b)

Table 11.6 Intake rate of tap water for children in Sweden

Age [years]	Average intake of tap water [L·d^{-1}]	95th percentile intake of tap water [L·d^{-1}]
4	0.131	0.375
8	0.152	0.152
11	0.179	0.575

Livsmedelsverket (2006)

Data from the Swedish Food Administration (Livsmedelsverket 2002) show that the WHO-values for adults correspond approximately to the 95th-percentile for water consumption. The average intake of coffee, tea and water in Sweden is 1.2 L/d for women, of which 0.6 L are tap-water. The 95th percentile was 2.4 L/d. For men, the average intake was 1 L/d, of which 0.3 L were tap water. The 95th percentile was 2 L/d. For children, the average intake of tap water and 95th percentiles (Livsmedelsverket 2006) are shown in Table 11.6. These data do not include the intake of water in the form of soup and other foodstuffs. Therefore, the WHO-value of 1 L/d is appropriate for a high, but realistic, consumption of drinking water for children.

11.4.3.2 Data for Volatilisation and Dermal Pathways

In order to calculate exposure due to volatilisation of contaminants in domestic water, parameters are needed for volatilisation (usually not scenario specific) and for the exposure (scenario specific). The volatilisation parameters are model specific and can include data for the mass transfer coefficients for the gas and water phase respectively and the molecular weight of the contaminant, shower water temperature and original water temperature, droplet size and falling time. The exposure parameters consist of the inhalation rate and the exposure time. Standard parameters are found in US EPA (1991), Brand et al. (2007) and OVAM (2004).

Data on exposure parameters and methods to estimate dermal uptake can be found in US EPA (2004b), Brand et al. (2007) and OVAM (2004).

11.4.4 Reliability and Limitations

Exposure estimates based on contaminant concentration in tap water are relatively reliable. Average consumption of drinking water is well known although there is a large variability between individuals. Also, the volatilisation from domestic water can generally be well described by models. The models are usually derived from observation of radon emissions from domestic waters for which there exists large amounts of data compared to what is available for ordinary soil contaminants. However, variations may be large in human behaviour (duration and frequency of showering, use of appliances for washing and dishing) and water consumption of household appliances. The models tend to be more reliable for the more volatile contaminants than for less volatile, but since it is for the former that this pathway is of greatest importance this may not be a major limitation.

There is a considerably larger degree of uncertainty associated with the estimation of how contaminated soil can contaminate groundwater and the resulting contaminant concentration in well water or tap water.

11.4.5 Verification and Validation

Several experimental studies have been performed to study inhalation exposure from showering (e.g., Moya et al. 1999; Xu and Weisel 2003) and from other household appliances, such as dishwashers (Howard-Reed et al. 1999). These studies show that measurements of the fraction of contaminant that is volatilised and the concentrations obtained in bathroom air are in good agreements with predictions. However, it is important to use appropriate physio-chemical parameters, for example Henry's Law constants for the correct temperature.

11.5 Exposure Through Inhalation of Vapours Outdoors

Vapours of volatile contaminants in soil may be released to the ambient air and thereby cause exposure to humans during outdoor activities (see Fig. 11.3 for a schematic representation).

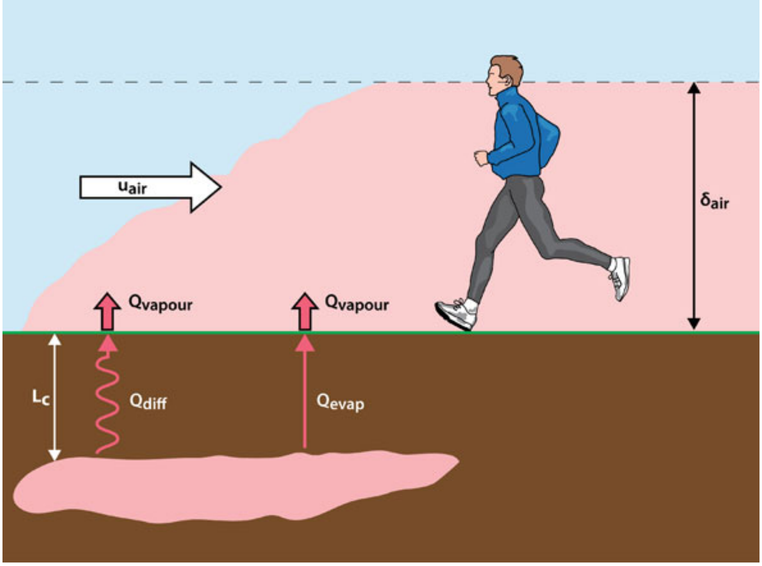


Fig. 11.3 Schematic representation of human exposure through vapours during outdoor activities

For sites with buildings, intrusion of vapours into buildings and the exposure due to inhalation of indoor air will usually be the dominating exposure pathway. This exposure pathway is treated in detail in McAlary et al. (Chapter 11 of this book). However, for sites without buildings and especially sites where large areas are contaminated with volatile contaminants, exposure due to inhalation of vapours in outdoor air may become of importance. In this section an overview is given of how the exposure due to inhalation of vapours in outdoor air can be calculated, including the calculation of outdoor air concentration and exposure.

11.5.1 Conceptual Model

Volatile soil contaminants will be distributed between soil particles, soil air and pore water. In the case of severe contamination also a free phase may be present. The transport of volatile contaminants through the soil can occur by diffusion or by advective transport. Diffusion in soil air will be the dominating process for volatile contaminants that to a large degree are present in soil air (contaminants with large Henry's Law constant). In soils with very high water saturation, diffusion of contaminants in vapour phase is very slow. However, dissolved contaminants may also diffuse in the pore water and may thus be transported towards the soil surface from where it can evaporate. This can be an important exposure pathway for contaminants with small Henry's Law constants. A common approximation is that the contaminant concentration in soil air and pore water is in equilibrium and that diffusion in the soil air and pore water can occur in parallel.

An advective flow can be caused by gas generation in the soil, thermal gradients in the soil or pressure differences in the atmosphere. In climates with a high evaporation rate, capillary rise can also cause an upward flow of water in the soil column, causing advective transport to the surface of contaminants in the soil water.

Vapours released from the soil surface will mix with the ambient air resulting in dilution. The degree of dilution will depend on wind speed, atmospheric stability and the roughness of the ground surface, but also on the size of the contaminated site and the height above ground at which human beings inhale contaminated air, i.e., at the height of the breathing zone.

11.5.2 Description of Models

The exposure due to inhalation of vapours in outdoor air is included in several Risk Assessment models. The starting point for the Risk Assessment may be the contaminant concentration in the soil, in the groundwater or in the soil air. Here we will focus on modelling approaches taking concentration in the soil as a starting point.

The mathematical model for exposure due to inhalation of vapours outdoors contains the following steps:

- calculation of outdoor air concentration;
- calculation of exposure.

11.5.2.1 Calculation of Outdoor Air Concentration

Calculation of outdoor air concentrations contains two parts, i.e., estimating the flux of volatile contaminant from the soil and estimating the dilution in the outdoor air. The transport of volatile contaminants in the soil is calculated in a way similar to that used when estimating indoor air concentrations, see Chapter 11 by McAlary et al., this book. Since the model for calculating the flux from the soil is closely interlinked to the calculation of dilution in outdoor air it will also be described in this section.

The different types of models used are based on the same principles, although the actual calculation methods may differ. Furthermore, the terminology may differ substantially between different models and may be quite confusing when a comparison of different models is made. Sometimes similar entities have different names and also similar names are used for different entities.

A commonly used approach (e.g., ASTM 2004; Environment Agency 2008; US EPA 1996a, 2002a) is to estimate a volatilisation factor relating the concentration in the ambient air with the concentration in the surface soil. Thus, the calculations include both the transport in the soil and the dilution in the ambient air. The contaminant is assumed to be located close to the ground surface, thus the loss of contaminant will lead to a reduction of the vapour release from the soil, with time. In order to handle this, the average vapour flux over the exposure duration is calculated either by assuming a diffusion driven vapour transport from an infinite source where the soil close to the surface is gradually depleted from contamination. This gives rise to a time-dependent vapour flux, where the volatilisation factor is estimated as the average flux over the chosen averaging time:

$$VF = \frac{\rho_s}{Q/C}\sqrt{\frac{4D_{\mathrm{eff}}}{\pi\tau}\frac{H}{K_{sw}\rho_s}} \tag{11.50}$$

in which

VF = the volatilisation factor [kg m^{-3}]
ρ_s = the dry bulk soil density [kg m^{-3}]
Q/C = the air dispersion factor [kg m^{-2} s^{-1} per kg m^{-3}]
D_{eff} = the effective diffusion coefficient in the soil [m^2 s^{-1}]
τ = the averaging time for surface emission vapour flux [s]
H = the dimensionless Henry's Law constant [–]
K_{sw} = the total soil-water portioning coefficient [m^3 kg^{-1}].

Alternatively, a constant release rate from a finite source is assumed to occur for the duration of the exposure, i.e. is the total amount of the contaminant is assumed to be volatilised during the exposure duration. In this case the volatilisation factor is given by:

$$VF = \frac{d_s \rho_s}{{}^{Q}\!/_{C} \cdot \tau_{\exp}} \tag{11.51}$$

in which

d_s = the thickness of the contaminated layer [m]
$\tau_{\exp}$ = the exposure duration [s].

The second approach is to calculate separately the vapour flux either due to diffusion, as is used in the CSOIL model from the Netherlands (Brand et al. 2007), or through the combined effect of diffusive transport and transport due to an evaporative flux, as in the Flemish model Vlier-Humaan, (OVAM 2004). In this case the volatile contaminant has to be located at some depth below the soil surface. Furthermore it is assumed that the source is infinite and that the soil concentration is not affected by evaporation. The diffusive flux through the soil is given by:

$$Q_{\mathrm{diff}} = \frac{D_{\mathrm{eff}}\left(C_{\mathrm{vap,soil}} - C_{\mathrm{vap,amb}}\right)}{L_c} \tag{11.52}$$

in which

Q_{diff} = the diffusive flux from the soil surface [kg·m^{-2}·s^{-1}]
$C_{\mathrm{vap,soil}}$ = the concentration in soil air [kg·m^{-3}]
$C_{\mathrm{vap,amb}}$ = the concentration in the ambient air above the soil [kg·m^{-3}]
L_c = the diffusion length in the soil [m].

This is simplified by assuming that the concentration in the soil air is much less than the concentration in the ambient air above the soil. Furthermore, the diffusivity is transformed to relate to the total concentration of contaminants in the soil, as follows:

$$Q_{\mathrm{diff}} = \frac{D_{\mathrm{soil}} C_{\mathrm{pore}}}{L_c} \cdot \frac{V_w}{P_w} \tag{11.53}$$

in which

D_{soil} = the total diffusivity in the soil [m^2 s^{-1}]
C_{pore} = the concentration in soil pore water [kg m^{-3}]
V_w = the volume fraction of water in the soil [–]
P_w = the mass fraction of contaminant in pore water [–].

In the case when the mass fraction in pore water is zero, the diffusivity is transformed to relate to the concentration in pore water or the total concentration of contaminants in the soil, as follows:

$$Q_{diff} = \frac{D_{soil} C_{soil} \rho_s}{L_c} \tag{11.54}$$

in which

C_{soil} = the total concentration in the soil [kg kg^{-1}]

In the Flemish model (Vlier-Humaan), the release due to the evaporative flux of water and due to diffusion over the boundary layer at the soil surface is also considered. The evaporative flux is given as:

$$Q_{evap} = C_{pore} E_v \tag{11.55}$$

in which

Q_{evap} = the evaporative flux from the soil surface [kg m^{-2} s^{-1}]
E_v = the evaporation rate [m^3 m^{-2} s^{-1}].

The diffusion over the boundary layer is given by:

$$Q_{vapour} = \frac{D_{eff}(C_{vap,surf} - C_{vap,amb})}{X_a} \tag{11.56}$$

in which

$C_{\text{vap,surf}}$ = the concentration in soil air at the soil surface [kg·m^{-3}]
X_a = the thickness of the boundary layer [m].

If the sum of the diffusive flux, Qdiff, and the evaporative flux, Qvap, is greater than the diffusion over the boundary layer, Qvapour, it is assumed that the boundary layer will limit the release of volatile contaminants into the ambient air. In other cases, the release will be determined by the sum of the diffusive flux, Qdiff, and the evaporative flux, Qvap.

In both approaches outlined in Eqs. (11.50) to (11.51) and Eqs. (11.52) to (11.56), respectively, the dilution in the ambient air is described by an air dispersion factor or a dilution velocity (Q/C). These are equivalent entities defined as the quotient of the mass flux per unit area and unit time from the soil and the concentration in the air at the receptor height. The air dispersion factor is often given in the units g·m^{-2}·s^{-1} per kg m^{-3}, while for the dilution velocity the units are shortened to cm·s^{-1} or m·h^{-1}.

Various methods are used to determine the dilution velocity ranging from:

- Assuming mixing in the air volume is defined by a specified wind speed and mixing height.
- Calculation of the mixing height from boundary layer theory.
- Using external atmospheric dispersion models to calculate site-specific and area specific dilution velocities.

The first type, i.e., assuming mixing in the air volume is defined by a specified wind speed and mixing height, is used in the RBCA model (ASTM 2004) where the dilution velocity is described by:

$$Q/C = \frac{U_{air}\delta_{air}}{W} \tag{11.57}$$

in which

U_{air} = the wind speed above ground in the ambient mixing zone [m s^{-1}]
δ_{air} = the ambient mixing zone height [m]
W = the dimension of the soil source zone parallel to the wind direction [m].

The ambient mixing zone is often taken to be the height of the receptor, i.e., the height above ground at which human beings inhale contaminated air, i.e., at the height of the breathing zone.

Several different methods are used to estimate the mixing height and the effective wind velocity for mixing. An approximation used by the Danish EPA (Miljøstyrelsen 2002) is that at low wind velocities the mixing height is proportional to the length of the contaminated area. The empirical relationship that the mixing height is 0.08 times the length of the contaminated area. Thus, Eq. (11.58) can be simplified as follows:

$$Q/C = 0.08U_{air} \tag{11.58}$$

This approach is also adopted in the Swedish model for Soil Quality Standards (Naturvårdsverket 2009).

A more complex model for determining the dilution velocity is used in the CSOIL and Vlier-Humaan models. The mean wind speed in the mixing layer is calculated from measured wind speeds at a height of 10 m assuming a logarithmic wind speed profile above the ground surface. The effect of obstacles on the ground is considered by using a *roughness length*. The height of the mixing layer is calculated as a function of atmospheric stability, roughness length and the length of the contaminated area.

For the UK CLEA model (Environmental Agency 2008) a different approach is used, in which air dispersion factors are tabled for thirteen cities in the UK for four different sizes of source areas (from 100 to 20,000 m^2) and two different receptor heights corresponding to adults and children. The tabled values are derived from calculations performed with the US EPA dispersion model AERMOD PRIME, calculating annual average air concentrations using local meteorological data.

11.5.2.2 Calculation of Exposure

The calculating of exposure to contaminants by inhalation of vapours outdoors differs, depending on the Toxicological Reference Value which is used for the contaminant being studied. There are two main approaches:

- For some contaminants, Toxicological Reference Values are available in the form of reference concentrations in air, RfC-values, in mg contaminant per m^3, or as an inhalation risk values, which can be used to give the concentration of contaminant in air equivalent to an acceptable risk.
- For other contaminants, Toxicological Reference Values are available only as reference doses, RfD-values, or as risk-based values for oral intake, in mg contaminant/$\mathrm{kg}_{\text{body weight}}$/d.

In the first case, exposure calculations therefore derive the contaminant concentration in air for comparison with the RfC-values. In the latter case, in some countries, the daily intake of contaminants via inhalation of vapours is calculated from the inhalation rate and the occupancy period and compared to RfD-values (see Chapter 10 by McAlary et al., this book).

11.5.3 Input Parameters

11.5.3.1 Diffusivities

The *diffusivity* of contaminants in the vapour phase in the air-filled soil pore volume is a function of water saturation. The *effective diffusivity* in the vapour phase ($D_{e,g}$) can be estimated according to Millington and Quirk (1961):

$$D_{e,g} = D_{0,g} \frac{a^{\frac{10}{3}}}{\varepsilon^2} \tag{11.59}$$

in which

$D_{0,g}$ = the diffusivity of the contaminant in free air [$\mathrm{m}^2 \cdot \mathrm{s}^{-1}$]
ε = the total porosity [m^3 pore space per m^3 bulk soil]
a = the soil air content [m^3 air filled pore space per m^3 bulk soil]

The effective diffusivity of dissolved contaminants in the water pore space ($D_{e,w}$) can be calculated in a similar way as:

$$D_{e,w} = D_{0,w} \frac{w^{\frac{10}{3}}}{\varepsilon^2} \tag{11.60}$$

in which

$D_{0,w}$ = the diffusivity of the contaminant in free water [$\mathrm{m}^2 \cdot \mathrm{s}^{-1}$]
w = the soil water content [m^3 water filled pore space per m^3 bulk soil]

The total effective diffusivity (calculated assuming equilibrium contaminant in air – contaminant in water) is then given by:

$$D_{eff} = D_{e,g} + \frac{D_{e,w}}{H} \tag{11.61}$$

in which

H = the dimensionless Henry's Law constant [–].

11.5.3.2 Meteorological Parameters

The estimation of the *dilution of contaminants in the ambient air* requires information on a number of meteorological parameters, primarily wind speed and mixing height. In order to calculate the mixing height, data concerning atmospheric stability and surface properties are needed. If an external atmospheric dispersion model is used, a number of site-specific meteorological data will be needed. However, external models are not likely to be used for specific sites, but rather to be used to derive typical regional air dispersion factors, as has been done in the US and the UK. Thus, the type of data needed for these models will not be discussed here.

Depending on the model used, wind speed may refer to different heights above the ground surface. The wind speed used in the RBCA model is the ambient air velocity in the mixing zone, using a default value of 2.25 m s^{-1}. The Danish model (Miljøstyrelsen 2002) also uses the air velocity in the mixing zone, but uses a lower wind velocity (0.1 m s^{-1}) corresponding to calm weather conditions. However, for contaminants where the Toxicological Reference Values are based on long-term effects, e.g., carcinogens, a higher wind velocity (1 m·s^{-1}) is used.

The Dutch and the Flemish models (CSOIL and Vlier-Humaan) use as input the wind speed at a height of 10 m, corresponding to the wind speeds given in meteorological statistics. The default value used in the models is 5 m·s^{-1}. This data is used in the model to calculate the wind speed at the height at which human beings inhale contaminated air, i.e., at the height of the breathing zone.

The RBCA model assumes a fixed mixing height (default value 2 m) related to the height of the breathing zone. The Danish model assumes that the mixing height is proportional to the length of the contaminated area and is therefore not included as input. In the Dutch and Flemish models the mixing height is calculated in the model based on data on atmospheric stability, surface roughness and the dimensions of the contaminated area.

11.5.3.3 Receptor Height

In both the RBCA and the Danish models a complete mixing is assumed within the mixing zone so receptor height is not used as a parameter. The Dutch and Flemish models use a height at which human beings inhale contaminated air, i.e., at the height of the breathing zone of 1.5 m for adults and 1 m for children. The UK model (CLEA) uses a height of the breathing zone for adults of 1.6 m and for children of 0.8 m.

11.5.4 Influence of Physical Properties

The soil properties are very important for determining the release of the contaminants from the soil. The diffusivity of vapours can differ several orders of magnitude between a dry soil and a wet soil. Soil porosity may also be important, but to a more limited extent. Also the depth of the contaminant in the soil, determining the length of the diffusion path, will be of importance for the release. The meteorological conditions, such as wind speed and atmospheric stability are very important for the resulting dilution.

11.5.5 Influence of Human Behaviour

Volatile contaminants are present in outdoor air due to their widespread use. Petrol stations, dry cleaners, and other commercial/industrial facilities are important sources of volatile contaminants to outdoor air. Indoor air may also become affected from outdoor air containing volatile contaminants.

The exposure times and also the height at which human beings inhale contaminated air, i.e., at the height of the breathing zone, will be determined by the type of human activities occurring at the site. For example exposure may be higher in playgrounds, picnic areas and parks. The type (and intensity) of human activity has an important effect on breathing rates and the volume of air inhalation.

11.5.6 Reliability and Limitations

It is generally recognized that modelling of exposure through outdoor air inhalation is relatively uncertain. The factors that cause uncertainty are explained in detail, in the following text.

Modelling of vapour release to outdoor air is based on the assumption that the different phases of the contaminant (vapour, dissolved and sorbed to soil) are in chemical equilibrium. This may not be the case, for example when transport in the soil is fast or when the contaminant is a constituent in free phase liquids, e.g., petrol, oil or chlorinated solvents. The assumption of equilibrium may give an overestimation of the release (e.g., vapour phase concentrations are often less than equilibrium concentrations due to the loss to the atmosphere), and, hence, of the human health risk involved. In some cases an underestimation may occur, e.g., when large parts of the pore water or the soil are not available to the contaminant. Furthermore, any degradation of the contaminant as it migrates through the soil or when it reached the ambient atmosphere is generally not considered, which may overestimate the concentrations of readily degradable contaminants.

The models for the inhalation of vapours outdoors aim at predicting an effective average value over a relatively long time scale. A number of the important input parameters such as water saturation of the soil and wind speed, however, will vary considerably on a very short timescale. Thus, the variation in both release and dilution will be substantial due to changing weather conditions. There may also be correlations between the parameters, for example hot dry periods give dryer soils with larger releases and often lower wind speeds. There may also be a correlation with the risk of exposure as people are more likely to stay outdoors during such periods.

Despite these limitations the models are likely to overpredict exposure in most situations if conservative data on soil moisture and wind speeds are used. The available models are therefore judged to be fit for the purposes of identifying when inhalation of outdoor air may be an important exposure pathway, i.e., in first tier Risk Assessment.

11.5.7 Verification and Validation

The models used for atmospheric dilution are all, to some extent, based on empirical data and can be expected to give reasonable approximations for the situations they are derived for, generally dispersion over relatively flat surfaces. The models are simplifications of more complex dispersion models that have been tested and validated. However, when extrapolating to more complex situations (e.g., forest or built-up areas) their ability to predict dilution will be more limited. In cases when this exposure pathway is judged to be critical, actual measurements of air concentrations can be made in site-specific Risk Assessments. However, as actual concentrations may vary considerably depending on the meteorological conditions, measurements over extended periods may be needed.

Measurements of outdoor air samples may also be a part of a general Risk Assessment for sites with volatile contaminants. A comparison of measurements in soil pore air, indoor air and outdoor air is recommended to get a better picture of the origin of the air contamination and for assessing the risks.

11.6 Exposure Through Inhalation of Dust

This exposure pathway relates to the inhalation of dust originating from contaminated soils. Dust inhalation can occur both indoors and outdoors, and both adults and children may be exposed. Exposure through ingestion of dust is described in Bierkens et al. (Chapter 6 of this book).

The exposure pathway through inhalation of dust considers long-term exposure levels, when dust arises principally from wind action on the soil surface. Short-term events with increased dust concentrations in air, for example when soils are mechanically disturbed during digging, are not considered in this chapter, though they may need to be considered when evaluating short-term or acute health risks.

There is a large amount of evidence that the contaminant concentration in dust is related to the total exposure to contaminants, see for example Thornton et al. (1990). Some, but not all, of the exposure to contaminated dusts is via the inhalation pathway.

People who inhale dust from a contaminated site can be exposed in two ways, i.e.,

- small particles will pass into the respiratory systems and may be deposited in the lungs;
- larger particles will get caught in the nasal passages and throat and eventually be swallowed.

In the latter case, i.e., larger particles that will get caught in the nasal passages and throat and eventually be swallowed, the uptake of contaminants and the toxicological response will be the same as when soil is directly ingested. In addition, the methods used to estimate soil intake, i.e., trace element excretion, cannot differentiate between swallowed dust and direct oral intake. Therefore, swallowed dust is included in the ingestion of soil exposure pathway (see Chapter 6 by Bierkens et al., this book). The remainder of this section will therefore exclusively treat inhaled dust that reaches the lungs.

11.6.1 Conceptual Model

People can be exposed by the inhalation of fine-grained material which is suspended from contaminated sites. This is, however, only one of many sources of dust. Dust particles in ambient air originate from a number of sources: traffic, combustion, long-range transport and local resuspension. Most measurements performed on air dust concentration focus on small size particles (with an aerodynamic diameter less than 10 μm, PM10). The finer fraction of dust in the air, PM10 or smaller, is usually considered to be the most relevant for assessing exposure by inhalation of dusts (Putaud et al. 2003).

Contaminant concentration in the soil's fine-grained fraction, which gives rise to fine-particulate dusts, can differ from the contaminant concentration in the fraction of soil which undergoes chemical analysis (usually <2 mm). Often, contaminant concentrations are higher in the fine fraction (Bright et al. 2006; Young et al. 2001).

Important parameters for exposure are:

- The concentration of particles from the resuspension of soil in the inhaled air.
- The fraction of soil particles deriving from the contaminated area.
- The fraction of dust particles which is respirable.
- Contaminant concentration in respirable particles compared with the contaminant concentration measured by chemical analyses of soil.
- Inhalation rate.
- Exposure time.
- Fraction of the dust which is exhaled directly and not retained in the lungs.

In addition, when exposure to contaminated dusts occurs indoors, the transport of dust from outdoors to indoors must be accounted for.

11.6.2 Mathematical Equations

11.6.2.1 Concentrations of Contaminants in Dust in Air

Two main approaches have been taken to quantifying the dust concentration in air:

- In the first approach, the dust concentration in air is specified, i.e., the amount of dust per cubic metre of air. The specified concentration is the long-term average concentration, and therefore a number of correction factors may be applied to measured dust concentrations to take into account variation with time and meteorological conditions. The fraction of the total amount of dust which is contaminated is then specified for the air where exposure takes place. When using this approach, a number of other factors are also used, such as the fraction of soil-derived dust which comes from the contaminated area and the fraction of dust which is respirable. These factors are normally given as parameters to the model and will be further treated in Section 11.6.3.
- In the second approach, algorithms are used to calculate a particulate emission factor from the soil and the lateral dispersion of the dust particles. This approach is used by for example the models RBCA and CLEA, though the algorithms used in the two models are different. These models calculate directly the emission of PM10 particles from contaminated land and the concentration of dust from the contaminated area in ambient air.

These two approaches are illustrated in Fig. 11.4.

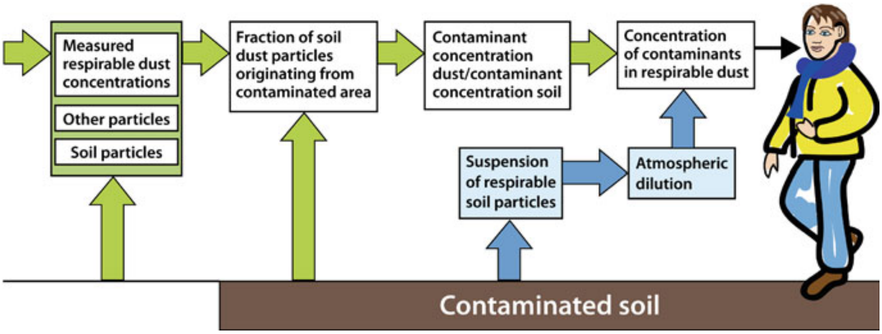


Fig. 11.4 Modelling approaches for quantifying exposure to contaminants through the inhalation of dust

11.6.2.2 Derivation of the Particulate Emission Factor

An example of derivation of particulate emission factors (PEF) is shown below. This expression is used by US EPA (1996b) for the calculation of Soil Screening Levels, and by the CLEA model (Environment Agency 2008).

$$PEF = Q/C \cdot \frac{1}{J_w} \tag{11.62}$$

in which

Q/C = the air dispersion factor or dilution velocity [kg m^{-2} s^{-1} per kg m^{-3}]
J_w = the PM_{10} emission flux, [kg m^{-2} s^{-1}].

The expression is based on the approach of Cowherd et al. (1985) to calculating wind erosion and includes two factors. The first factor (*Q/C*) describes the ratio between particle emission flux and the concentration of dust in air at receptor height. The same factor is also used for calculating the dilution of vapours released to ambient air, and has been treated in Section 11.5.2. Generic values are suggested for the air dispersion factor by both US EPA and CLEA, with values depending on geographic location (and thereby on meteorological conditions), on the size of the contaminated site and on the height of the receptor. The generic values are based on the results of dispersion models.

The second term, $1/J_w$, describes the inverse of the PM10 emission flux from the soil surface. J_w is based on a semi-empirical approach described by Cowherd et al. (1985), as follows:

$$J_w = 0.036 \cdot (1 - V) \cdot (U_m/U_t)^3 \cdot F(x) \cdot \frac{1}{3600} \tag{11.63}$$

in which

V = the fraction of outdoor surface cover [–]
U_m = the mean annual wind speed at height 10 m [m·s^{-1}]
U_t = the threshold value of wind speed at height 10 m [m·s^{-1}]
$F(x)$ = an empirical function of x [dimensionless].

An explanation of how to obtain site-specific estimates of V, U_m, U_t and $F(x)$ can be found in Cowherd et al. (1985) and is also given in the technical background of the CLEA model (Environmental Agency 2008).

Other models are available for determining the dilution velocity used in calculating the particulate emission factors. These models are discussed in Section 11.5 of this chapter on inhalation of vapours outdoors.

11.6.2.3 Calculation of Exposure

The calculation of exposure follows the same methodology as for calculating the exposure due to inhalation of vapours, see Section 11.5.2.2.

For contaminants where Toxicological Reference Values are available in the form of reference concentrations in air, RfC-values (in mg/m^3), the derived contaminant concentration in air is compared with the RfC-value. For contaminants where Toxicological Reference Values are available only as reference doses, RfD-values (in mg/kg$_{\text{body weight}}$/d), the daily intake of contaminants via inhalation of dusts may be calculated from the inhalation rate and the occupancy period and compared with the RfD value.

There has been some discussion about whether the extrapolation of the toxicity of a substance via the oral intake route to the inhalation intake route is valid. The method relies on the assumption that the route of intake is irrelevant to the dose delivered to the organ where effects occur. This assumption is not supported by the available data, and is particularly irrelevant for substances inducing local effects and for which route-to-route extrapolation has been found to be unreliable (Rennen et al. 2003), Extrapolation from the oral route to the inhalation route has been suggested to be unreliable, especially for less volatile contaminants (US EPA 1996b; Wilschut et al. 1998). There are limited data available on which to base conclusions about route-to-route extrapolation. It is not clear what factors, beside bioavailability are responsible for the difference in toxicity between routes.

Generally, RfC-values are unavailable for the less-volatile substances. If exposure by inhalation (e.g. of dusts) occurs for these substances, route-to route extrapolation may be the only means of accounting for the exposure pathway. In future, the problem of route to route exposure may be avoided by developments in the use of pharmokinetic models, and the expression of contaminant dose in mg/g tissue in target organs, regardless of exposure pathway.

In some cases adjustments are made to account for limited availability of a contaminant for uptake. This is normally done by applying a bioavailability factor. This relates the availability for uptake in the lungs of a contaminant in respired dust to the availability of the contaminant used when determining the Toxicological Reference Values. However, methods for testing the bioavailability of respired particles are still under development and not yet standardised and furthermore, the actual bioavailability of the reference contaminant used for setting the Toxicological Reference Values is often not very well known. Therefore, in most models a bioavailability factor of 1.0 is used.

11.6.3 Input Parameters

11.6.3.1 Dust Concentrations in Air

The dust concentration in air varies widely between different sites and also over time. In the Swedish countryside, for example, an annual average PM10-concentration is 8–16 μg/m^3 (Socialstyrelsen 2005). In urban environments, the background concentration is 14–20 μg/m^3, but can be much higher in streets with heavy traffic, in many places exceeding the Environmental Quality Standards of a maximum annual average value of 40 μg/m^3.

In the Dutch CSOIL model (Brand et al. 2007; Van den Berg 1991/1994/1995) the applied dust loading factors are 53 and 70 μg/m^3 for indoor and outdoor air, respectively. This is based on Hawley (1985), who further assumes that the concentration of suspended dust in the air inside homes is 75% of that in the outside air. Paustenbach et al. (1997) recommended an indoor dust level of 50 μg/m^3 for generic models. Oomen and Lijzen (2004) reviewed studies carried out with personal sampling devices, and concluded that a dust concentration of 60 μg/m^3 is probably representative of moderately crowded places such as homes, whereas a value of 100 μg/m^3 is probably more representative of crowded places such as classrooms. In a study by Janssen et al. (1998), outdoor concentrations (mean PM10 of 41.5 μg/m^3) exceeded indoor concentrations (mean PM10 of 35.0 μg/m^3), but these levels underestimated adult exposure measured with personal devices (mean PM10 of 61.7 μg/m^3). For children, mean indoor PM10 levels in classrooms (81–157 μg/m^3) were higher than outdoor levels (30 μg/m^3). The personal samples showed the same levels as the classrooms (mean value 105 μg/m^3). The studies reviewed by Oomen and Lijzen (2004) showed that indoor dust levels are dependent on the number of people in a room and on the activity levels, therefore dust levels are likely to be higher when children are present. Dust levels were also dependent on height above floor level (Beamer et al. 2002), and children are exposed to higher amounts of PM10 than adults in the same room.

11.6.3.2 Fraction of Dust from the Contaminated Site

The fraction of material with mineralogic origin in dust is between 6 and 25% (Putaud et al. 2003) of the total amount of dust. Resuspended particles can account for about half of the total weight (Hedberg et al. 2001), though a large fraction of these particles may be resuspended from road surfaces.

In the Swedish model for generic Soil Quality Standards, the annual average value for dust concentration from a contaminated area in outdoor air is estimated to be 5 μg/m^3 based on an urban background PM10 concentration of 20 μg/m^3, by assuming that up to half of the PM10 concentration has a mineralogic origin, and that 50% of the material of mineralogic origin is derived from the contaminated area. The resulting concentration is assumed to be reasonably conservative. This assumption is supported by model calculations of PM10 emissions from soil at three places in Sweden (RIVM-VITO 2006).

11.6.3.3 Concentration of Dust in Indoor Air

The concentration of dust in indoor air is assumed to be 75% of the outdoor concentration, based on the fact that the PM2.5 fraction (which is about 75% of the PM10 size fraction) is the fine fraction which is transported to indoor air.

Oomen and Lijzen (2004) reviewed estimates of the relative contribution of exterior soil to house dust. Estimates from different articles ranged from > 80% to 8%, depending on a wide variety of site-specific factors and methodological approaches. The US EPA (2002b) uses a default value for the contribution of soil to indoor dust of 70% in the IEUBK (Integrated Exposure Uptake BioKinetic) model which is used to predict the risk of elevated blood lead levels in children who are exposed to environmental lead from many sources. The value of 70% is based on empirical relationships between soil and dust lead concentrations. Oomen and Lijzen (2004) suggest that house dust can contain between 30 and 70% soil, and that estimates in this range are appropriate for the assessment of exposure to soil contaminants via house dusts.

11.6.3.4 Fraction of Dust Which is Respirable

The dust concentration in air which is specified in the first approach is usually based on measurements of particle concentrations in air. Measurements are often made of the concentration of the particle fraction PM10, i.e., particles with the same aerodynamic properties as a sphere with diameter up to 10 μm. The PM10 fraction is a measure of particles that can pass into the lungs. In some cases, a finer fraction, PM2.5 is also measured. This fraction is equivalent to the fine particles which can reach the alveoli. The PM2.5 fraction is usually about 50–90% of the PM10 fraction, with a mean of 73% (Putaud et al. 2003).

In the Dutch CSOIL model, the retention of particles reaching the lungs is assumed to be 75% for both adults and children (Brand et al. 2007; Van den Berg 1991/1994/1995), which is equivalent to the fraction of PM2.5 particles in the PM10 size fraction. Contaminants present in PM10 particles may also have adverse effects, even if they do not reach the alveoli.

11.6.3.5 Contamination in Dust

Contamination levels are usually higher in fine particles than in soil in average. However, these relationships vary depending on soil type and the origin of the contamination (Bright et al. 2006). Normally, soil investigations are not performed on extremely fine particles in the soils. Young et al. (2001) measured lead concentrations in soils and dust (PM10) and found that the concentration in the PM10 fraction was up to 8 times higher than in the soil on average. This is consistent with measurement of fine particle materials found on samples from Swedish wood treatment plants and glass works. In the Swedish model an enrichment factor of 5 is used for calculating the generic soil guidelines. In the German UMS model an enrichment factor of 4 is used for inorganic contaminants and a factor of 8 for organic contaminants (UMS 1997).

11.6.4 Inhaled Volume

Assessments of the risk from exposure through inhalation of contaminated soil particles is usually based on estimates of the volume of air inhaled over a longer period of time. The volume of air inhaled varies with physical characteristics, such as gender, age, body size and fitness as well as on work rate and activity level. The inhalation rate shows large short-term variations, depending on activity levels. The assumed volume of air inhaled over a longer period of time takes into account the short term variations.

The volume of air which is inhaled per day is assumed to be 20 m^3/d for adults in a number of Risk Assessment models, e.g., the Dutch model (CSOIL), and the

Table 11.7 Recommended inhalation volumes for long-term exposure [m^3 day^{-1}]

ECETOC 2001			US EPA 2006		
Age class [years]	Female	Male	Age [years]	Female	Male
<1	4.5				
1–2	6.8		1	8.8	8.5
3–5	8.3		5	12.2	12.7
6–8	10				
9–11	13	14	10	12.4	13.4
12–14	12	15	15	13.4	15.3
15–18	12	17	17	14.8	19.4

EU model for evaluating contaminants ECB (2003). These values are based mainly on data for ICRP's Reference Man (ICRP 1975) and data from US EPA (1997). ECETOC (2001) use data from the USA (US EPA 1997). ICRP (1975) give mean values of 23 m^3 per day for men and 21 m^3 per day for women. The US data for metabolically consistent breathing rates (Layton 1993) is based on three energy-based approaches for estimating daily inhalation rates: average daily intakes of food energy from dietary surveys, (with an upward adjustment to account for under reporting of food consumption); average daily energy expenditure based upon the ratio of total daily expenditure to basal metabolism; and daily energy expenditures based upon time-activity data. Values from all approaches were below those of the ICRP reference values and the mean values for adults between the ages of 19 and 65 years were 15.2 m^3/d for males and 11.3 m^3/d for females. In a study by Lordo et al. (2006) estimates of the inhalation volume were calculated as a function of oxygen consumption (derived from energy expenditure). For eighteen year olds, the estimates were 16.4 m^3/d for males and 12.0 m^3/d for females.

US EPA (2006) recommends age-specific inhalation rates for children, also based on the data of Lordo et al. (2006). ECETOC (2001) also gives age-specific inhalation rates for children. These rates are shown in Table 11.7. CLEA's recommended inhalation rates according to age and sex for residential and commercial land use are based on Lordo et al. (2006). In other models, the choice of default value is determined partly by the age-group being considered. In CSOIL, a value of 7.6 m^3/d is used for children with a body weight of 15 kg and a rate of 20 m^3/d for adults.

11.6.5 Influence of Soil Properties

The soil particle size distribution determines how much wind is needed before dust is generated at a site, therefore soil texture is important with respect to the amount of dust in air. The surface cover of a site of course is important in determining how much dust can be generated. The surface may be covered by vegetation, with a range of densities, or by made-surfaces (e.g., asphalt, concrete) or buildings. Other non-erodable surfaces, e.g., large stones or rocks, may also be important.

Meteorological conditions, especially the intensity and frequency of wind, affect both the emission and dispersion of particulate matter. The amount and distribution of rainfall is also important, as rainfall directly affects the dust content of the air as well as determining the soil moisture content, which is important for dust generation. Surface erosion by water may also be important in that the redistribution of the fine-fraction of soil materials may affect subsequent wind erosion.

Chemical and physical properties of the soil affect the bioavailability of contaminants in soil particles for uptake in the lungs.

11.6.6 Influence of Human Behaviour

Important factors are occupancy on the contaminated site and time spent indoors versus outdoors. Since human activity will vary considerably, normally, standardised values are used in order to obtain conservative estimates of the exposure.

The type of activity on the site will also affect exposure, e.g., extensive gardening work or sport activities may increase the concentrations of contaminated dust in the air and also increase the inhalation rate.

11.6.7 Reliability and Limitations

As the specified dust concentrations in air are based on measurements, the values should be of the correct order of magnitude. However, the dust concentration in air varies with time over a large interval and is largely incident-steered, for example dry windy conditions with high rates of dust generation, or wet, still conditions with low rates of dust generation. Therefore, it is difficult to derive average dust concentrations from measurements.

A further uncertainty is the fraction of respirable dust which is contaminated, which depends on the dilution of dust from the contaminated area after suspension from the soil surface with other dust. Local conditions are important for this parameter, particularly downstream of the site of dust generation to the place where exposure takes place.

The concentration of contaminants in the respirable fraction of dust is an important uncertainty in the model. This parameter has a direct effect on the calculated exposure for contaminants, but there is relatively little information on the relationship between the contaminant concentration in the fine fraction of dust and the contaminant concentration in the soil as a whole, or in different particle-size fractions of the soil.

When no reference air concentration (RfC value) is available for a contaminant, data on the oral toxicity of the contaminants may be extrapolated to the inhalation pathway in some cases in order to assess the risks from the inhalation pathway. This extrapolation gives rise to uncertainties associated with the differences in uptake, metabolism and toxicity between the two exposure pathways. These uncertainties have not been quantified, but may be important, particularly for contaminants which have a direct effect on the function of the lungs.

11.7 Exposure Through Dermal Uptake

11.7.1 Significance

At contaminated sites contaminants may come in contact with the skin either in surface water and groundwater or in soils and sediment.

Dermal uptake has often been omitted from Exposure Assessments based on the assumption that dermal uptake is a minor contributor to total exposures. In many cases that assumption is correct. The relative contribution of dermal uptake to total exposures to contaminants in soil varies depending on the exposure scenarios being considered and the nature of the contaminants being assessed. Volatile contaminants are typically assumed to have minimal potential for dermal uptake due to the likelihood that they will volatilize faster than dermal uptake could occur. Most metals are also expected to have limited dermal uptake due to a low tendency to partition into skin lipids. This leaves semi-volatile organic contaminants as the focus of most dermal uptake assessments.

Typical residential exposures for young children are likely to include the potential for substantial dermal contact with soil through routine play activities, whereas adults may only have substantial soil contact with skin while engaged in selected activities such as gardening. Workers may also have highly variable soil contact, with construction workers or utility workers digging trenches having the greatest potential exposure via the dermal route. As intake assumptions are varied within any one scenario, the relative contributions from different exposure routes may also change. For example, if soil ingestion rates in young children are reduced, the relative contribution of dermal uptake to total exposures would increase. Similarly, if a larger area of skin is assumed to be in contact with soil, the importance of dermal uptake will increase.

11.7.2 Conceptual Model

Systemic absorption of contaminants from the skin surface requires a series of transfers; contaminants must cross into the stratum corneum (outer skin layer) followed by transfer into the epidermis (live skin cell layer), and then into the blood. The rate and extent of these transfers is a function of many factors, including the condition and thickness of the skin and the characteristics of the contaminant itself. Skin thickness varies over different areas of the body, so the nature of the exposed body parts should be specified when estimating dermal uptake. Lipophilic contaminants such as PCBs, some PAHs, dioxins, and some pesticides typically have greater potential for dermal uptake compared with ionic or hydrophilic contaminants. In the case of soil and sediment, the rate and extent of dermal uptake will also be controlled by the nature of the soil particles that adhere to skin and by the interactions of the contaminant with the soil particles.

The model typically used to estimate dermal uptake of contaminants from soil and sediment is derived from the idea that a finite dose of contaminant comes in contact with the skin during each exposure event, and that some fraction of that applied dose is subsequently absorbed (US EPA 2004b). This model differs from the model used to estimate uptake from water, which is based on use of contaminant-specific permeability constants to estimate rates of absorption over a specified exposure period. Practical application of the soil absorption model and the most critical input parameters are described in the following subsections.

Due to the variable contribution of dermal uptake to exposures, initial screening evaluations may be useful to determine if there is a need to more fully characterize dermal uptake. Where dermal exposures are predicted to contribute 5% to 10 % or less to total exposures, it may not be necessary to include dermal uptake calculation in Risk Assessments.

Assessing human health risks from dermal uptake differs from assessments for other exposure routes, because there is little data from studies that directly measure systemic toxicity of environmental contaminants via dermal exposure. Typically, it is necessary to rely on oral Toxicological Reference Values to assess risks from dermal exposures. Most oral Toxicological Reference Values are stated in terms of intake ("external exposure") rather than uptake ("internal exposure"). Consequently, it is necessary to convert oral toxicological reference to an absorbed dose or uptake before using those values to assess dermal exposures.

For non-cancer effects, this conversion is accomplished by multiplying the oral Toxicological Reference Value by the fraction of the oral dose that is absorbed into the systemic circulation. For example, if the tolerable daily intake (TDI) is 100 $mg/kg_{body\ weight}/d$, and 5% of the dose is absorbed, then the tolerable absorbed dose will be 5 $mg/kg_{body\ weight}/d$ (i.e., 0.05 times 100). Conversion of a cancer slope factor to an absorbed dose basis requires dividing the oral slope factor by the fraction of dose absorbed. US EPA (2004b) provides a more detailed explanation of this process and recommended assumptions for the oral absorption of a limited number of contaminants.

11.7.3 Mathematical Equations

Dermal uptake for chronic exposure is typically calculated in a two step procedure. In the first step, the absorbed dose for a single exposure event is calculated from the soil concentration, dermal absorption fraction and assumed amount of soil adhering to the skin. The calculation of the absorbed dose per event is as follows:

$$DA\text{event} = C\text{soil} \times CF \times AF \times ABSd \tag{11.64}$$

Where:

$DA\text{event}$ = Absorbed dose per event (mg/cm^2-event)
$C\text{soil}$ = Concentration of contaminant in soil (mg/kg_{dw})
CF = Unit conversion factor ($1 \text{ x } 10^{-6}$ kg/mg)
AF = Adherence factor (mg/cm^2-event)
$ABSd$ = Dermal absorption fraction, contaminant-specific (unitless).

In the second step, the average daily dose over the exposure period is calculated by incorporating estimates of the number of events per day, exposure frequency in days per year, exposure duration in years, and average body surface area exposed. For exposures that span childhood into adulthood, it may also be necessary to prorate and combine the average daily doses during childhood with those during adulthood. The calculation of the average daily dose over the exposure period is as follows:

$$DADsoil = \frac{DAevent \times EF \times ED \times EV \times SA}{AT \times BW} \tag{11.65}$$

Where:

DADsoil = Dermal absorbed dose (mg/kg-day)
DAevent = Absorbed dose per event (mg/cm^2-event)
EF = Exposure frequency (days/year)
ED = Exposure duration (years)
EV =Event frequency (event/d)
SA = Skin surface area available for contact (cm^2)
AT = Averaging time (days)
BW = Body weight (kg).

11.7.4 Input Parameters

Critical input parameters for estimating dermal uptake from soil include dermal absorption fractions, skin adherence factors, and skin surface area exposed.

11.7.4.1 Dermal Absorption Fractions

As described above, many factors affect the absorption of contaminants through the skin. For studies of dermal absorption of contaminants in soil many aspects of study design affect the relevance of the results in predicting behaviour of contaminants in soils at contaminated sites. The principal problem with studies reported in the literature is the failure of most studies to use weathered soils from sites. Most laboratory studies have been conducted using soil to which a contaminant dissolved in a solvent has been added. In some studies the solvent was not even evaporated before the soil-contaminant mixture was applied to the skin. Contaminant mobility and bioavailability in soil have been shown to decline over time (Alexander 2000), so for that reason much empirical data on dermal absorption from soil is likely to overestimate the contribution of dermal exposures to contaminated site exposures. A related study design flaw is use of contaminant concentrations exceeding the soil saturation limit. Such studies are not truly studies of absorption from soil (Spalt et al. 2008).

Other critical design factors in dermal absorption studies are particle size fractions and soil loading. Typically only particles less than 150 μm in diameter are expected to adhere to skin, and for dry soils a more reasonable cutoff may be 65 μm or less (Choate et al. 2006). Consequently, studies will be more representative of expected exposure conditions if only the fine particle fraction is tested. Characterizing the particle size distribution is also a critical step in planning the soil loading to be used in a study. When the soil loading exceeds a monolayer of particles on the skin, it has been shown that the absorption fraction declines. Therefore, studies conducted using high soil loadings may underestimate dermal absorption at soil loadings more typical of those on people at contaminated sites. Spalt et al. (2008) described many of these issues in a recent critical review, and offered design recommendations for site-specific studies.

US EPA (2004b) provided default dermal absorption fractions for ten compounds, including arsenic (0.03), cadmium (0.001), chlordane (0.04), 2,4-dichlorophenoxyacetic acid (0.05), DDT (0.03), TCDD and other dioxins (0.03 or if soil organic content exceeds 10%, then 0.001), lindane (0.04), benzo(a)pyrene and other PAHs (0.13), PCBs (0.14), and pentachlorophenol (0.25). None of these default values are based on studies that used weathered soils representative of contaminated sites. The differences, however, can be substantial. For example, a recent study of dermal absorption of arsenic from weathered soils found that in contrast to US EPA's default value of 0.03, absorption was negligible (Lowney et al. 2007). The US EPA (2004b) recommended default value for other semi-volatile contaminants is 0.01. According to US EPA (2004b) volatile contaminants and other inorganic contaminants do not need to be evaluated for dermal absorption.

11.7.4.2 Soil Adherence Factors

The adherence factor describes the amount of soil that sticks to the skin. It is expressed in terms of soil mass per unit surface area (mg/cm^2). Soil adherence to skin is influenced by the properties of the soil, the part of the body exposed, and the activities of the exposed individual. Empirical data on adherence factors is summarized by US EPA (2004b). Soil adherence studies have examined the adherence of soil to various body parts of children and adults engaged in a variety of activities. Adherence factors recommended by US EPA (2004b) are weighted averages based on body part- and activity-specific adherence factors. Updated recommendations and more detailed documentation are provided in the *Child-Specific Exposure Factors Handbook* (US EPA 2008) and the *Exposure Factors Handbook 2009 Update* (US EPA 2009).

For a child resident, US EPA (2004b) recommended default soil adherence factors of 0.04 and 0.2 mg/cm^2 for central tendency estimates and reasonable maximum exposure cases, respectively. The central tendency value is a geometric mean derived for daycare children playing indoors and outdoors and for children playing in dry soil. The reasonable maximum exposure value is a geometric mean derived for children playing in wet soil.

For an adult resident, US EPA (2004b) recommended default soil adherence factors of 0.01 and 0.07 mg/cm^2 for central tendency estimates and reasonable maximum exposure cases, respectively. The central tendency value is a geometric mean derived from a study of groundskeepers, and the reasonable maximum exposure value is a geometric mean derived for gardeners.

As summarized in US EPA (2004b), geometric mean adherence factors developed for pipe layers working in dry soil, construction workers, heavy equipment operators, utility workers, and pipe layers working in wet soil are 0.07, 0.1, 0.2, 0.2, and 0.6 mg/cm^2, respectively (US EPA 2004b). For typical adult commercial/industrial workers, US EPA (2004b) recommended default soil adherence factors of 0.02 and 0.2 mg/cm^2 for central tendency estimates and reasonable maximum exposure cases, respectively. The central tendency value is a geometric mean derived for commercial groundskeepers, and the reasonable maximum exposure value is a geometric mean derived for heavy equipment operators and utility workers.

US EPA (2004b) recommended that dermal exposure to sediments be evaluated in a manner similar to dermal exposure to soils, with two modifications to account for differences between the two media. First, primarily sediments that are exposed (i.e., not underwater) for at least a portion of the year should be evaluated, because sediments covered by water are likely to be washed off the skin before the individual reaches shore. Second, sediment sample results must be adjusted to reflect wet weight so they are consistent with the reporting methods of soil adherence studies. Wet soil sticks to the skin more easily than dry soil, so adherence factors for wet soil are higher than for dry soil. Geometric mean adherence factors developed for children playing in wet soil, pipe layers working in wet soil, and reed gatherers are 0.2, 0.6, and 0.3 mg/cm^2, respectively (US EPA 2004b).

Shoaf et al. (2005a, b) investigated sediment adherence for adults clamming and children playing in tide flats. Weighted median and 95% upper confidence interval (UCI) loading rates for hands, forearms, lower legs, and feet combined reported in these papers are 0.35 and 0.87 mg/cm^2, respectively, for adults and 4.8 and 7.0 mg/cm^2, respectively, for children. The soils at both of these sites were very sandy, resulting in approximate monolayer loadings (which is defined as the complete coverage of the skin with one layer of particles) of 13 and 31 mg/cm^2 for the adult and child scenarios, respectively.

11.7.4.3 Skin Surface Area

Dermal exposure to contaminants is affected by the surface area of skin being in contact with the impacted media. US EPA (2004b) provided recommendations for skin surface area that may have contact with exposure media for residents and workers. Updated recommendations and more detailed documentation are provided in the *Child-Specific Exposure Factors Handbook* (US EPA 2008) and the *Exposure Factors Handbook 2009 Update* (US EPA 2009). The recommended body surface area estimates are based on 50th percentile values for specific body parts for the U.S. population from the national health and nutrition examination survey (NHANES II, see Centers for Disease Control and Prevention 2006) and are intended to correlate with average body weight values used in child and adult exposure estimates. The estimates for skin surface area are based on population mean values for males and females combined.

For typical residential exposures, US EPA (2004a, b) recommended using skin surface areas of 2,800 and 5,700 square centimetres (cm^2) for child and adult residents, respectively, assuming the exposed skin surface is limited to the head, hands, lower legs, and forearms (short-sleeved shirts, shorts, and shoes are assumed to be worn). Exposure scenarios for worker contact with soil assume that skin contact is limited to the head, hands, and forearms, which have a skin surface area of 3,300 cm^2 (US EPA 2004b). Estimates are provided of skin surface area for individual body parts so that area estimates for additional scenarios may be calculated.

11.7.5 Reliability and Limitations

It is generally recognized that exposure estimates for dermal uptake are highly uncertain. Few studies have been conducted that accurately characterize the potential for dermal absorption under the conditions present at contaminated sites. Volatile contaminants can be excluded from dermal uptake assessments with a high degree of confidence, and most metals can also likely be excluded without any substantial concern for underestimating total exposures. That leaves semi-volatile contaminants as the critical group to evaluate. Available empirical data in the group of semi-volatile contaminants suggest that dermal uptake could be substantial enough to require quantification in most Risk Assessments. However, dermal uptake from weathered soils may be much less than suggested by studies in which the contaminants are added to soils shortly before testing.

Reliability could be increased when oral Toxicological Reference Values are converted to an absorbed dose or uptake before using those values to assess dermal exposures. Moreover, applying realistic values for skin surface, for example depending on normal clothing characteristics in regard to climatic conditions, could also improve the quality of site-specific dermal exposure calculations.

References

Alexander M (2000) Aging, bioavailability, and overestimation of risk from environmental pollutants. Environ Sci Technol 34:4259–4265

Andelman JB (1990) Total exposure to volatile organic compounds in potable water. In: Significance and treatment of volatile organic compounds in water supplies, pp 485–504

ASTM (2004) Standard guide for risk-based corrective action, Standard E 2081-00 (Reapproved 2004), ASTM International, American society for testing and materials, USA

Beamer P, Castano A, Leckie JO (2002) Vertical profile particulate matter measurements in a California daycare. Proceedings: Indoor Air 2002, pp 103–108

Brand E, Otte PF, Lijzen JPA (2007) CSOIL 2007: an exposure model for human risk assessment of soil contamination. RIVM report 711701054 RIVM, Bilthoven, the Netherlands

Bright DA, Richardson GM, Dodd M (2006) Do current standards of practice in Canada measure what is relevant to human exposure at contaminated sites? I: a discussion of soil particle size and contaminant portioning in soil. Human Ecol Risk Assess 12(3):591–605

California Department of Toxic Substances Control (2009) http://www.dtsc.ca.gov/AssessingRisk/ctox_model.cfm. Retrieved 20 Dec 2009

Carlon C, Swartjes F (2007) Analysis of variability and reasons of differences. In: Carlon C (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities, JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

Centers for Disease Control and Prevention (2006) National health and nutrition examination survey data. National Center for Health Statistics (NCHS). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Hyattsville, MD. http://www.cdc.gov/nchs/nhanes.htm. Retrieved 20 Dec 2009

Choate LM, Ranville JF, Bunge AL, Macaalday DL (2006) Dermally adhered soil: 1. Amount and particle size distribution. Integr Environ Assess Manag 2:375–384

Cowherd C, Muleski GE, Englehart PJ, Gillette DA (1985) Rapid assessment of exposure to particulate emissions from surface contamination sites. Report EPA/600/8–85/002. United States Environmental Protection Agency, Washington, DC

DEFRA and EA (2002) The contaminated land exposure assessment (CLEA) model: technical basis and algorithms. DEFRA (Department for Environment, Food and Rural Affairs) and EA (The Environment Agency), EA R&D publication CLR 10, Bristol, UK

Dowdy D, McKone T, Hsieh D (1996) Prediction of chemical biotransfer of organic chemicals from cattle diet into beef and milk using the molecular connectivity index. Environ Sci Technol 30(3): 984–989

ECB (2003) Technical guidance document on risk assessment. Part I, risk assessment for human health. EUR 20418/EN/1. European Chemicals Bureau

ECETOC (2001) Exposure factors sourcebook for European populations (with focus on UK data), Technical report no. 79, European centre for ecotoxicology and toxicology of chemicals, Brussels, June 2001

Environment Agency (2008) Updated technical background to the CLEA model. Science report – SC050021/SR2. Environment Agency, UK

ExpoFacts (2006) Exposure Factors Sourcebook for Europe. http://expofacts.jrc.ec.europa.eu

European Commission (2008) The European Union System for the Evaluation of Substances (EUSES 2.1). Available via the European Chemicals Bureau, http://ecb.jrc.it/Euses

Foster SA, Chrostowski PC (1986) Integrated household exposure model for use of tap water contaminated with volatile organic chemicals. Proc 79th Ann Meeting of Air Pollution Control Association, Minneapolis

Gawronski SW (2000) Toxic metals accumulation, distribution, and diversity of tolerance mechanisms in higher plants. Warsaw Agricultural University, Intercost Workshop on Bioremediation-Sorrento 2000

GCNC (2002) Impact of chemical discharges, vol. 2, annexes 3 and 4: available information on agronomical parameters for the radioecological working group (in French). Groupe Chimique Nord-Cotentin, http://www.irsn.fr/FR/base_de_connaissances/Environnement/surveillance-environnement/GRNC/Pages/Radioecologie-nord-cotentin.aspx. Reterived 23 August 2010

Gesell TF, Pichard HM (1978) The contribution to radon in tap water to radon indoor concentrations. Presented at the 3rd international symposium on natural radiation, Huston, TX

GRNC (2002) Analyses of sensitivity and uncertainty for risk of leukaemia related to Nord Cotentin’s nuclear installations, Annex 1: Probabilistic distributions of input parameters (in French). Groupe Radioécologique Nord-Cotentin, http://www.irsn.fr/FR/base_de_connaissances/Environnement/surveillance-environnement/GRNC/Pages/Radioecologie-nord-cotentin.aspx. Reterived 23 August 2010

Hawley JK (1985) Assessment of health risk from exposure to contaminated soil. Risk Anal 5: 289–302

Hedberg E, Hansson H-C, Johansson C, Vesely V, Wideqvist U, Kristensson A (2001) Characterization of particles in Lycksele and Gothenburg. ITM air pollution laboratory, Stockholm University department of nuclear physics, Lund University May 2001. ITM Rapport 92

Hess CT, Weiffenbach CV, Norton SA (1982) Variations of airborne Rn-222 in houses in Maine. Environ Int 8:59–62

Howard-Reed C, Corsi RL, Moya J (1999) Mass transfer of volatile organic compounds from drinking water to indoor air: the role of residential dishwashers. Environ Sci Technol 33: 2266–2277

IAEA (1994) Handbook of parameter values for the prediction of radionuclide transfer in temperate environments. International Atomic Energy Agency, Technical reports series no. 364

IAEA (2001) Generic models for use in assessing the impact of discharges of radioactive substances to the environment. International Atomic Energy Agency, Safety reports series no. 19

ICRP (1975) Reference Man; anatomical, physiological and metabolic characteristics, ICRP Publication 23. International Commission on Radiation Protection. Pergamon, New York

INERIS (2003) Evaluation of the impact of atmospheric deposition of a large combustion plant using coal on human health (in French), INERIS (National Institute of the Industrial Environment and Risks) Final report DRC-03–45956 / ERSA-no. 92-RBn, France

INERIS (2009) Set of equations for modelling exposure from contaminated soil contamination or due to industrial emissions (in French), INERIS (National Institute of the Industrial Environment and Risks) DRC-08-94882-16675A, France

INSEE (1991) Consumption and places of purchase of foodstuffs in 1991 (in French), Bertrand M, INSEE (National Institute of Statistics and Economical studies)

Intawongse M, Dean JR (2006) In-vitro testing for assessing oral bioaccessibility of trace metals in soil and food samples. TrAC Trends Anal Chem 25(9):876–886

Janssen NAH, Hook G, Brunekreef B, Harssema H, Mensink I, Zuidhof A (1998) Personal sampling of particles in adults: relation among personal, indoor and outdoor air concentrations. Am J Epidemiol 147:537–547

Layton DW (1993) Metabolically consistent breathing rates for use in dose assessments. Health Phys 64:23–36

Lowney YW, Wester RC, Schoof RA, Cushing CA, Ruby MV (2007) Dermal absorption of arsenic from soils as measured in the Rhesus monkey. Toxicol Sci 100:381–392

Lordo B, Sanford J, Mohnson M (2006) Revision of the metabolically-derived ventilation rates within the exposure factors handbook. Batelle Institute, Columbus OH. Prepared for US EPA/ORD

Lijzen JPA, Baars AJ, Otte PF, Rikken MGJ, Swartjes FA, Verbruggen EMJ, Van Wezel AP (2001) Technical evaluation of the Intervention Values for Soil/sediment and groundwater. RIVM report 711701023, February 2001. RIVM, Bilthoven, the Netherlands.

Little JC (1992) Applying the two-resistance theory to contaminant volatilization in showers. Environ Sci Technol 26(7):1341–1349

Livsmedelsverket (2002) Riksmaten 1997–98, Food habits and nutrition in Sweden. Method and analysis of results (in Swedish), Becker W, Pearson M (eds) National Food administration, Sweden

Livsmedelsverket (2006) Riksmaten – children 2003. Food habits and nutrition in children in Sweden (in Swedish), Enghardt Barbieri H, Becker W, Pearson M (eds) National Food Administration, Sweden

McLachlan M (1994) Model of the fate of hydrophobic contaminants in cows. Environ Sci Technol 28(13):2407–2414

Miljøstyrelsen (2002) Guidelines on remediation of contaminated sites. Environmental Guidelines No. 7 2002, Vejledning fra Miljøstyrelsen. Miljøstyrelsen, Danmark

Millington RJ, Quirk JM (1961) Permeability of porous solids. Trans Faraday Soc 57:1200–1207

Moya J, Howard-Reed C, Corsi RL (1999) Volatilization of chemicals to indoor air from contaminated water used for showering. Environ Sci Technol 33:2321–2327

Naturvårdsverket (2009) Guideline values for contaminated land. Model description and guidance (in Swedish). Swedish Environmental Protection Agency, Stockholm

Oomen AG, Lijzen JPA (2004) Relevancy of human exposure via house dust to the contaminants lead and asbestos. RIVM report 711701037/2004, National institute for public health and the environment, Bilthoven, the Netherlands.

OVAM (2004) Basic information for risk analysis, Part 3H, Formulae for Vlier-Humaan (in Dutch)

Paustenbach DJ, Finley BL, Long TF (1997) The critical role of house dust in understanding the hazards posed by contaminated soils. Int J Toxicol 16:339–362

Putaud J-P et al (2003) A European aerosol phenomenology, physical and chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe, European Commission, Report nr. EUR 20411 EN, 2003

Rennen MA, Bouwman T, Wilschut A, Bessems JGM, De Heer C (2003) Oral-to-inhalation route extrapolation: a critical assessment. Regul Toxicol Pharmacol 39(1):5–11

RIVM-VITO (2006) De Raeymaecker B, Cornelis C, Provoost J, Joris I, De Ridder K, Lefebre F, Ottte P Lijzen J, Swartjes F (2007) Evaluation of the Swedish guideline values for contaminated sites – cadmium and polycyclic aromatic hydrocarbons, VITO/RIVM, 2006/IMS/R/. VITO, Belgium

RTI (2005) Methodology for predicting cattle biotransfer factors. Research Triangle Institute, prepared for US EPA, Office of Solid Waste, Research Triangle Park (NC), USA

Shoaf MB, Shirai JH, Kedan G, Schaum J, Kissel JC (2005a) Adult dermal sediment loads following clam digging in tide flats. Soil Sediments Contamin 14:463–470

Shoaf MB, Shirai JH, Kedan G, Schaum J, Kissel JC (2005b) Child dermal sediment loads following play in a tide flat. J Exp Anal Environ Epid 15:407–412

Simonich SL, Hites RA (1995) Organic pollutant accumulation in vegetation. Env Sci Tech 29:2905–2914

Socialstyrelsen (2005) Environmental health report 2005. National board of health and welfare, Institute of Environmental Medicine, Stockholm county council. ISBN 91-7201-931-X. (in Swedish)

Spalt EW, Kissel JC, Shirai JH, Bunge AL (2008) Dermal absorption of environmental contaminants from soil and sediment: a critical review. J Exp Sci Env Epid 19:119–148

Swartjes FA (2007) Insight into the variation in calculated human exposure to soil contaminants using seven different European models. Integrated Environ Assess Manage 3(3):322–332

Swartjes FA, Dirven-Van Breemen EM, Otte PF, Van Beelen P, Rikken MGJ, Tuinstra J, Spijker J, Lijzen JPA (2007) Towards a protocol for the site-specific human health risk assessment for consumption of vegetables from contaminated sites. RIVM report 711701040. RIVM, Bilthoven, the Netherlands

Thornton I, Davies DJA, Watt JM, Quinn MJ (1990) Lead exposure in young children from dust and soil in the United Kingdom. Environ Health Perspect 89:55–60

Travis C, Arms A (1988) Bioconcentration of organics in beef, milk and vegetation. Environ Sci Technol 22:271–274

UMS, Hempfling R, Doetsch P (1997) Scientific support and development of a human health exposure model for historical contamination. Final report (in German), ARGE Fresenius-focon. F&E Vorhaben 109 01 215, im Auftrag des Umweltbundesamtes, Berlin

The University of California, McKone T (1993) CalTOX, 1993, a multimedia total exposure model for hazardous-waste sites, UCRL-CR-111456, Part I–IV, Lawrence Livermore National Laboratory, Livermore

US DOE (2003) Environmental transport input parameters for the biosphere model. Office of civilian radioactive waste management, Wasiolek M, US Department of Energy, ANL-MGR-MD-000007 REV 01

US DOE (2004) Office of civilian radioactive waste management, Rasmuson K, Agricultural and environmental input parameters for the biosphere model. US Department of Energy, ANL-MGR-MD-000006 Rev02

US EPA (1991) Risk assessment guidance for superfund (RAGS) vol 1. Human health evaluation manual, part B, development of risk-based preliminary remediation goals, EPA/540/H-92/003, Publ 9285-7-01B, US Environmental Protection Agency, Washington, DC

US EPA (1996a) Soil screening guidance, User's guide EPA540/R-96/018 and technical background document EPA/540/R-95/128, US Environmental Protection Agency, Washington, DC

US EPA (1996b) Soil screening guidance: technical background document, 2nd edn. EPA/540/R95/128, US Environmental Protection Agency, Washington, DC

US EPA (1997) Exposure factors handbook revised, PB98-124217, US Environmental protection agency, National center for environmental assessment, Office of research and development, US Environmental Protection Agency, Washington, DC

US EPA (2002a) Supplemental guidance for developing soil screening levels for superfund sites. OSWER 9355.4-24, US Environmental Protection Agency, Washington, DC

US EPA (2002b) Overview of the IEUBK model for lead in children, EPA report number PB 99-9635-8. OSWER, US Environmental Protection Agency, Washington, DC

US EPA (2004a) Exposure and human health reassessment of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and related compounds, part I: estimating exposure to Dioxin-like compounds, vol 3: site-specific assessment procedures, US EPA (US environmental protection agency), National centre for environmental assessment, Office of research and development, Washington, DC

US EPA (2004b) Risk assessment guidance for superfund volume i: human health evaluation manual (part e, supplemental guidance for dermal risk assessment), EPA/540/R/99/005, OSWER 9285.7-02EP, PB99-963312, US Environmental Protection Agency, Washington, DC

US EPA (2005) Office of solid waste. Human health risk assessment protocol for hazardous waste combustion facilities (HHRAP). US environmental protection agency, report EPA/530/R-05–006, US Environmental Protection Agency, Washington, DC

US EPA (2006) Child specific exposure factors handbook (external review draft), EPA/600/R/06/096A, US Environmental Protection Agency, Washington, DC

US EPA (2008) Child specific exposure factors handbook, EPA/600/R06/096F, National center for environmental assessment, Office of research and development, US Environmental Protection Agency, Washington, DC

US EPA (2009) Exposure factors handbook 2009 update, External review draft, PA/600/R09/052A, National center for environmental assessment, Office of research and development, US Environmental Protection Agency, Washington, DC

Van den Berg R (1991/1994/1995) Human exposure to soil contamination. A qualitative and quantitative analyses resulting in proposals for human health C screening values, Modified version of the original report Van den Berg 1991/Van den Berg 1994. RIVM report 725201011, RIVM, Bilthoven, the Netherlands

Veerkamp W, Ten Berge W (1994) The concepts of HESP. Reference manual, human exposure to soil pollutants,Version 2.1a, Shell International Petroleum, The Hague

WHO (1993) Guidelines for drinking water quality, 2nd edn, vol 1. Recommendations

WHO (2006) Guidelines for drinking water quality, 3rd edn. Incorporating first addendum, vol 1, Recommendations, 2006. WHO, Geneva

Wilschut A, Houben GF, Hakkert BC (1998) Evaluation of route-to-route extrapolation in health risk assessment for dermal and respiratory exposure to chemicals, TNO report V97.520. TNO Nutrition and Food Research Institute, Zeist, the Netherlands

Wu FH, Zhang G, Bachir DML (2004) Differences in growth and yield in response to cadmium toxicity in cotton genotypes. J Plant Nutr Soil Sci 167(1):85–90

Xu X, Weisel C (2003) Inhalation exposure to haloacetic acids and haloketones during showering. Environ Sci Technol 37:569–576

Young TM, Heeraman DA, Sirin G, Ashbaugh L (2001) Resuspension of contaminated soil as a source of airborne lead. Research division air resources board, Sacramento, CA, Final Project Report Contract Number 97-325, 31 August 2001

Chapter 12
Hazard Assessment and Contaminated Sites

Andrew Langley

Abstract *Hazard Assessment* is a key component of Human Health Risk Assessment and is comprised of the steps of *Hazard Identification* and *Dose-response Assessment.* Hazard Identification examines the capacity of a contaminant to cause adverse health effects in humans and other animals using data from a range of toxicological and epidemiological sources. Dose-response Assessment considers both qualitative and quantitative toxicological and epidemiological information to estimate the incidence of adverse effects occurring in human populations at different exposure levels. The conclusions from *Hazard Assessment* are assessed with those from *Exposure Assessment* to enable *Risk Characterisation*. This chapter provides an overview of the toxicological and epidemiological tools used for *Hazard Assessment* to enable a broader understanding of the Risk Assessment of contaminated sites and the principles underlying the development of risk-based policy and Soil Quality Standards.

Contents

A. Langley (✉)
Population Health Queensland, Maroochydore, QLD, Australia
e-mail: Andrew_Langley@health.qld.gov.au

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_12,

Abbreviations

ADI	Acceptable Daily Intake (WHO)
BMD	Benchmark Dose
DOH	Department of Health (United Kingdom)
DNA	Deoxyribonucleic acid
IARC	International Agency for Research on Cancer
IPCS	International Programme on Chemical Safety
LOAEL	Lowest Observed Adverse Effect Level
Mg/kg$_{\text{bw}}$/d	mg/kgbodyweight/day
MTD	Maximum Tolerated Dose
NHMRC	National Health and Medical Research Council (Australia)
NOAEL	No Observed Adverse Effect Level
NOEL	No Observed Effect Level
OECD	Organisation for Economic Co-operation and Development
PCB	Polychlorinated biphenyl
PM$_{10}$	Particulate Matter 10 μ

PTWI	Provisional Tolerable Weekly Intake (WHO)
Q_1*	The 95% upper confidence limit of the slope estimate used for the linear multi-stage model
R*f* D	Reference Dose (US EPA)
TDI	Tolerable Daily Intake (WHO)
US EPA	United States Environmental Protection Agency
WHO	World Health Organization

12.1 Hazard Assessment and Contaminated Sites

Hazard Assessment is a key component of Human Health Risk Assessment and is comprised of the steps of *Hazard Identification* and *Dose-response Assessment*.

The relationships between the components of Human Health Risk Assessment are detailed in Fig. 12.1, which also demonstrates the overlap with Risk Management.

Hazard Identification involves determining:

- what types of (adverse) health effects might be caused by the contaminant; and
- how quickly the adverse health effects might be experienced and their duration (Health Canada 1999).

The data for Hazard Identification will come from a range of toxicological and epidemiological sources. The toxicological sources may include in vitro, animal,

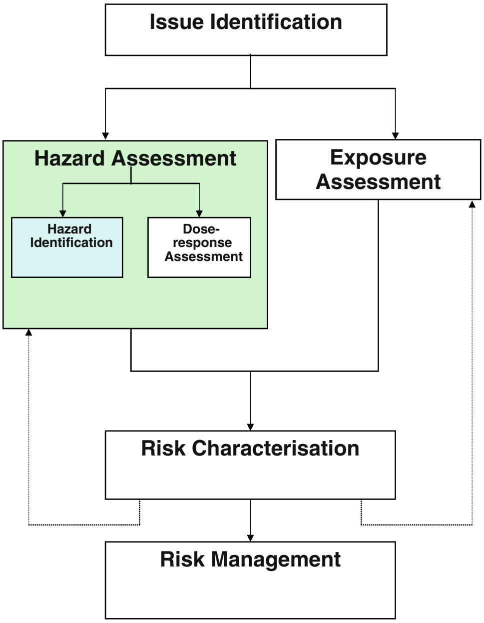


Fig. 12.1 Relationships of human health Risk Assessment and risk management

human and mechanistic studies. Not only may the contaminant need to be assessed but, in the case of contaminants, the breakdown products such as the four metabolites of atrazine (desethylatrazine, desisopropylatrazine, diamonochlorotriazine and hydroxyatrazine) when monitoring atrazine contamination of water.

Dose-response Assessment is the second component of Hazard Assessment and considers both qualitative and quantitative toxicity information to determine "the incidence of adverse effects occurring in humans at different exposure levels" (US EPA 1989, p. 1.6). Where available, human and animal evidence will be assessed as part of this process. Risk Assessment cannot be done without good dose-response information. Whereas constant doses can be used in animal studies, human exposures may be variable over time. This may be a significant source of uncertainty.

Although directed at ambient air standards setting, the Australian Ambient Air Quality Standards Setting: An Approach to Health-Based Hazard Assessment (NHMRC 2006) provides much useful information on Hazard Identification, assessing potential carcinogens and integrating animal and human data and the assessment of association and causation and complements the Australian Environmental Health Risk Assessment Guidelines.

For application of Toxicological reference values *with regard to contaminated site management*, see Chapter 5 by Swartjes and Cornelis, this book.

12.2 Hazard Identification

Hazard Identification examines the capacity of a contaminant to cause adverse health effects in humans and other animals (US EPA 1995). It is a qualitative description based on the type and quality of the data, complementary information (e.g., structure-activity analysis, genetic toxicity, pharmacokinetic), and the weight of evidence from these various sources (ibid). Key issues include (ibid):

- nature, reliability and consistency of human and animal studies;
- the availability of information about the mechanistic basis for activity; and
- the relevance of the animal studies to humans.

The Dose-response Assessment examines the quantitative relationships between exposure and the effects of concern. "The determination of whether there is a hazard is often dependent on whether a dose-response relationship is present" (ibid , p. 7). Important issues include:

- the relationship between the extrapolation models selected and available information on biological mechanisms;
- how appropriate data sets were selected from those that show the range of possible potencies both in laboratory animals and humans;
- the basis for selecting interspecies scaling factors to account for scaling doses from experimental animals to humans;

- relevance of the exposure routes used in the studies to a particular assessment and the interrelationships of potential effects from different exposure routes;
- environmental conditions (pH, organic matter, clay content, temperature);
- the relevance to the assessment of the expected duration of exposure and the exposure durations in the studies forming the basis of the dose-response assessment; and
- the potential for differing susceptibilities in population subgroups (ibid , p. 7).

Both qualitative and quantitative toxicity information is evaluated in assessing "the incidence of adverse effects occurring in humans at different exposure levels" (US EPA 1989, p 1.6).

Hazard Identification uses:

- Animal data. This is usually assessed by toxicological methods.
- Human data. This is usually assessed by epidemiological methods when groups of people are involved, or by toxicological methods when using case studies and acute chamber studies.
- Other data. This includes data such as structure-activity data or *in vitro* data assessed by toxicologists.

The data may come from a range of sources such as anecdotal data, case-report data and data collected from epidemiological registries (such as cancer or pregnancy outcome data). In each instance the quality of the study design and methodology and the resulting data will need to be rigorously assessed.

12.3 Hazard Identification-Toxicology

This chapter on toxicological evaluation is based in part on the draft OECD Monograph, Guidance Notes for the Analysis and Evaluation of Repeat-Dose Toxicity Studies, prepared for the OECD by the Chemicals Unit, Department of Health and Aged Care, Canberra, Australia, in cooperation with the US EPA and the Canadian Pest Management Regulatory Agency (PMRA). It was published as: Organisation for Economic Co-operation and Development OECD Series on Pesticides No. 10 *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies* (2000). OECD, Paris.

This chapter focuses on chemical hazards and in particular on some of the problems and pitfalls which may arise during an assessment of possible contaminant-related changes in parameters measured in toxicology studies conducted on a contaminant.

Toxicology studies have been designed to permit determination of toxic effects associated with exposure to chemical hazards. Such studies can provide information relating to toxic effects and potential health hazards likely to arise from single or repeated exposures, in terms of predicting potentially important toxicity end points and identifying potential target organs or systems. It is important to note that, over time, the scientific community will gain a better understanding of the mechanisms of toxicity and this may lead to changes in both methodology and interpretation of hazard data. Analysis and evaluation of toxicity studies should reflect scientific consensus at the time the data are reviewed.

12.3.1 Toxicity Testing – Major In Vivo Study Types

Hazard Identification mostly relies on the results of in vivo toxicity studies conducted according to standard protocols, e.g., OECD Test Guidelines (OECD 1998). The following types of studies are defined:

Acute toxicity studies are studies which investigate the effects of single doses of a contaminant. The LD_{50} test, or medium lethal dose test which records gross toxicity and mortality data over a 14 day post-dosing period, has been commonly employed, but newer tests ("limit" tests and "up-and-down" dosing methods) are now favoured as they reduce the numbers of animals required and reduce the suffering seen in the classical LD_{50} test.

The standard acute toxicity studies include tests for acute oral, dermal and inhalational toxicity, eye irritation, skin irritation and skin sensitisation. Such studies may serve as the basis for classification and labelling of a particular contaminant or contaminant mixture, and serve as an initial guide to possible toxic modes of action and in establishing a dosing regimen in sub-chronic toxicity studies.

Short term repeat-dose toxicity studies. A short-term study has been defined (WHO 1990) as "having a duration lasting up to 10% of the animal's lifespan, 90 days in rats and mice, or 1 year in dogs", although the US EPA considers a 1-year dog study to be a chronic study. The main purpose of short term repeat-dose testing is to identify any target organs and to establish dose levels for chronic exposure studies.

Chronic toxicity studies, or long-term studies, are defined as studies lasting for the greater part of the lifespan of the test animals, usually 18 months in mice, two years in rats (WHO 1987; 1990). The protocol for these studies may cover the investigation of chronic toxicity or carcinogenicity, or both.

Reproductive toxicity studies are studies designed to provide general information about the effects of a test contaminant on reproductive performance in both male and female animals, such as effects on mating behaviour, gonadal function, oestrous cycling, conception, implantation, parturition, lactation, weaning and neonatal mortality. These studies may also provide some information about developmental or teratogenic effects of the test contaminant. The conduct of and the results from these studies are very important to assess with care, since the reproductive process is critical for perpetuation of the species and factors or contaminant that alter or disrupt this process can have devastating consequences, both in fact and in public perception (Korach 1998).

Developmental toxicity studies are studies which examine the spectrum of possible in utero outcomes for the conceptus, including death, malformations, functional deficits and developmental delays (Tyl and Marr 1997). Exposure during sensitive periods may alter normal development resulting in immediate effects, or may subsequently compromise normal physiological or behavioural functioning later in life. Since some developmental processes can occur peri- or postnatally, protocols for developmental studies are being reviewed with the possibility of extending the dosing period in developmental toxicity studies from the period covering major organogenesis to cover the perinatal and early postnatal period.

Genotoxicity studies are designed to determine whether test contaminants can perturb genetic material to cause gene or chromosomal mutations. A large number of assay systems, especially in vitro systems, have been devised to detect the genotoxic or mutagenic potential of contaminant. Most authorities consider that a minimum set of data is required to define a mutagen/non-mutagen. These data usually consist of gene mutations in bacteria and mammalian cells and in vitro and in vivo cytogenetics. Newer assays which could provide additional information include the Comet assay, mutations in transgenic animals, fluorescent in situ hybridisation and cell transformation (IARC 1999).

12.3.2 Important Issues in Toxicity Testing and Assessment

12.3.2.1 Study Protocol and Design

Dosing Regimen

The purpose of toxicity studies is the detection of valid biological evidence for any toxic and/or oncogenic potential of the contaminant being investigated. Therefore, protocols should maximise the sensitivity of the test without significantly altering the accuracy and interpretability of the biological data obtained. The dose regimen has an extremely important bearing on these two critical elements.

Since the determination of dose responses for any observed effects is one of the objectives of repeat-dose studies, at least three dose levels are normally required, as well as controls. US EPA guidelines allow a limit dose of 1,000 mg/kg in chronic and sub-chronic studies. If this dose produces no observable toxic effects and if toxicity is not expected, based upon data on structurally-related contaminants, then a full study using three dose levels might not be considered necessary. Ideally, the dose selection should maximise the detection of potential dose-response relationships and facilitate the extrapolation of these to potential hazards for other species including humans. The largest administered dose should not compromise biological interpretability of the observed responses. For example, it is generally considered that the upper dose should not:

- cause a body weight decrement from concurrent control values of greater than 10–12%;
- in a dietary study, exceed 5% of the total diet because of potential nutritional imbalances caused at higher levels;
- produce severe toxic, pharmacological or physiological effects that might shorten duration of the study or otherwise compromise the study results;
- in a carcinogenicity study, alter survival in a significant manner due to effects other than tumour production.

High doses which overwhelm normal mechanisms for metabolism, detoxification and/or excretion, or produce severe tissue damage (i.e., necrosis, demyelination) can make interpretation difficult or lead to inappropriate conclusions about the extent of the hazard.

It is commonly accepted that the lowest dose should not produce any evidence of toxicity (i.e., allows the establishment of an NOAEL).

Dosing Route

For repeat-dose studies, the most convenient route of administration is by dietary admixture. However, depending on the possible route of exposure of the public or occupationally exposed workers to a contaminant or an environmental contaminant, it may need to be investigated by the dermal and/or inhalational route.

For dermal exposure the material, in a suitable vehicle, is applied to the clipped skin of rats, rabbits or guinea-pigs; OECD test guideline No. 410 recommend even application to an area representing about 10% of the total body surface area using methods that avoid dislodgement or evaporation.

The surface area of the respiratory tract is approximately 50–100 m^2 in the normal adult and much more air (about 5000 times, by volume) is inhaled each day than food or water is ingested (McClellan and Henderson 1989). Thus, exposure to airborne material is potentially greater than via dermal or oral exposure. Airborne material can be gases or vapours, liquid droplets or solutions, aerosols (solid and vapour components), or dry fibres or powders. As a consequence, the mechanisms needed to deliver contaminants to a test chamber in a form that can be inhaled are quite complex, particularly when coupled with the need to include measuring devices which can establish particle size, concentration and form of the material in the exposure chamber. Furthermore, many factors can influence the inhalation, deposition and retention of inhaled materials in the respiratory tract. Therefore, the conduct of inhalational studies is considerably more complex than equivalent studies by the dietary or dermal routes. Of critical importance, in both the conduct and assessment of such studies, in the need to establish what portion of the material delivered to the exposure chamber was in a respirable form.

Study Findings – Physiological, Pharmacological, or Toxic?

Responses produced by contaminants in humans and experimental animals may differ according to the quantity of the contaminant received and the duration and frequency of exposure e.g., responses to acute exposures (a single exposure or multiple exposures occurring within twenty four hours or less) may be different from those produced by sub-chronic and chronic exposures. Not all observed responses within a study, irrespective of exposure duration or frequency, will represent toxicity per se. They may encompass a range of effects from physiological through pharmacological and toxic manifestations. Although it sometimes may be difficult to make a clear distinction between these responses, an attempt to do so should be made. It is essential that all relevant toxicity end points (statistically and/or biologically significant) be identified for consideration when evaluating data for the presence or absence of non-toxic levels.

Physiological responses vary within limits which are in accord with the normal functioning of a living organism; examples of such response are the usual respiratory and pulse rate increases associated with increased physical activity, systemic changes associated with normal pregnancy, and those associated with homeostatic mechanisms. These variable factors are not important toxicity end points in sub-chronic and chronic exposure studies unless their fluctuations are abnormally altered by a dose regimen. If such alterations occur at a particular dose or are part of a dose-response relationship, they should be correlated with other toxicity end points which may be present.

Pharmacological responses are altered physiological functions arising from interaction of a contaminant with a cellular receptor site, are reversible, and are usually of limited duration following removal of the stimulus. Whilst some of these responses may be undesirable under certain circumstances, they are distinguished from toxic (adverse) responses by generally not causing injury.

Toxic responses may be reversible or irreversible, but are distinguished from other types of responses by being injurious and therefore adverse and harmful to living organisms or tissues. A contaminant which causes a physiological or pharmacological effect may produce a toxic response if the exposure is prolonged and/or if the dose is increased beyond a certain level. The reversibility or otherwise of such responses may also depend on these two factors. The reversibility or irreversibility of a histopathological change will depend on the ability of the injured organ or tissue to regenerate. For example, liver has a relatively great ability to regenerate and many types of injury to this organ are reversible. By contrast, differentiated cells of the central nervous system are not replaced and many injuries to the CNS are irreversible.

12.3.3 Assessment of the Quality of the Data Characterising the Hazard

The following considerations address the acceptability of experimental studies and the documentation provided:

- The adequacy of the experimental design and other experimental parameters, including: the appropriateness of the observational and experimental methods;

frequency and duration of exposure; appropriateness of the species, strain, sex and age of the animals used; the numbers of animals used per dosage group; justification of dose, route and frequency of dosing; and the conditions under which the contaminant was tested.

- There are many guidelines to the generation of scientifically valid data which concern good experimental design, laboratory practice and reporting, e.g., OECD and US EPA guidelines, and accepted codes of Good Laboratory Practice (OECD 1982; US EPA 1983).
- The competency and completeness of the conduct and reporting of the study.
- The effects of modifying factors that may result in major inequalities between control and test animals.

This qualitative consideration has more to do with the evaluation and interpretation of data than with acceptability of documentation. It is placed here because determination of the factors which may have a major influence on toxicological data needs to be made prior to the analysis of the data. There are many factors influencing the responses of experimental animals to experimental treatment; some of these are discussed by Doull (1980). Some influences may be quite subtle, as exemplified by studies performed by Thompson et al. (1982), in which it was noted that the onset of acute pulmonary oedema in rats being used in immune hypersensitivity studies was sudden and seasonal. Circadian rhythms and seasonal physiological variations can subtly influence experimental results. Such factors influencing animal responses can be troublesome when their effects are confused with or misinterpreted as toxic responses to treatment

The acceptability of reports and other technical information is primarily a scientific judgment. Therefore, the rationale for rejecting a Hazard Assessment study should be succinctly stated in the evaluation document.

12.3.4 Analysis and Evaluation of Toxicity Studies

Useful guidance documents for evaluating data and conducting assessments include the IPCS Environmental Health Criteria (EHC) monographs specifically EHC 6, 70, 104 and 141 (WHO 1978, 1987, 1990, 1992).

12.3.5 Analysis and Evaluation of Major Study Parameters

Concurrent control groups should always be used; notwithstanding the value of historical control ranges in tumorigenicity studies. It is generally not appropriate to rely on statistical comparisons with historical controls since the incidence of spontaneous lesions can vary significantly over time (and even between concurrent randomised control groups). Controls must be age-matched because some forms of toxicity represent no more than acceleration and/or enhancement of age-related changes.

The use of non-treated and vehicle-control groups aids assessment of effects due to vehicle or excipients. When a vehicle is used to deliver the doses of the contaminant understudy (e.g., a lipophilic contaminant delivered in corn oil), the need for vehicle-treated controls is paramount. Some parameters can be affected by animal handling (e.g., the serum enzyme ALT was raised in mice which were grasped around the body compared with unhandled or tail-handled mice; Swaim et al. 1985), so control animals should be treated in the same way as test animals.

Control animals must receive as much attention during the analysis and evaluation process as do the treated ones as any untreated animal or group may exhibit some signs of abnormality or drift from the norm for that species or strain.

Historical control data may be useful when evaluating the acceptability of the "normal" data obtained from control groups (Haseman et al. 1984; Paynter 1984).

12.3.5.1 Mortality/ Survival

The separation of deaths caused by factors unrelated to exposure to the test contaminant (e.g., acute or chronic infections, age or disease-related degenerative processes, anatomical abnormalities, negligent handling or accident) from toxicity-induced deaths is important. All data relating to moribund or dead animals during their study life, as well as the results of post-mortem examinations, should be scrutinised in an attempt to make this distinction. The US EPA guidelines state that the highest dose used in sub-chronic studies with non-rodents should not produce an incidence of fatalities which would prevent meaningful evaluation.

Changes in the protocol during the course of a study can complicate the analysis e.g., alterations in dosage levels can produce a confusing mortality pattern.

Any unusual mortality pattern should be explained by the test laboratory on biological or toxicological grounds. If overall mortality is high (i.e., significantly greater than expected for the particular colony and strain) for any repeat-dose study, or for a particular group within a study, a credible explanation should be provided.

Deaths which are clustered at a specific time period may reflect a spontaneous epidemic disease situation of limited duration. High mortality associated with infectious contaminant in treated groups, in the absence of such evidence in the concurrent control group, could indicate an immunosuppressive action of the contaminant being tested.

The effect of dietary intake on mortality needs to be considered. A contaminant administered in the diet may make the laboratory chow more or less palatable, may have a pharmacological stimulant or depressant effect on appetite, or may affect the partitioning of the nutrients in the food. Likewise, decreased water consumption (e.g., in the case of an unpalatable contaminant administered in the water) will lead to reduced food consumption. These effects may significantly influence longevity since it has been clearly shown in animal species that long-term dietary restriction very significantly increases lifespan (e.g., Tucker 1979). Conversely, excessive ad libitum intake of highly nutritious diets can reduce lifespan compared with the expected average lifespan for an animal species/strain.

12.3.5.2 Clinical Observations

Adverse clinical signs (gross observations) noted during the exposure period may correlate with toxicity end points or disease processes. These can be used as supportive evidence for dose-response relationships and may play a role in the determination of the NOEL/NOAEL. However, not all adverse clinical signs will correlate with pathological or morphological changes in organs or tissues. Some will be caused by biochemical or physiological effects, i.e., incoordination, muscle twitching, tremor, or diarrhoea may indicate acetylcholinesterase inhibition without any morphological changes being evident in nervous tissue.

Clinical observations such as palpable tumours or those which might be associated with neoplasia (e.g., haematuria, abdominal distension, or impaired respiration) may be useful in defining the time a tumour was first suspected as being present. Such signs might aid in the evaluation of decreased tumour latency in long-term rodent studies. They may also aid in determining cause of death. A statement of the correlations, or the lack thereof, between clinical signs and specific toxicity end points should be made in the evaluation.

The revised OECD test guidelines for 90-day oral toxicity studies in rodents and non-rodents (Test Guidelines 408 and 409) have placed additional emphasis on neurological end-points.

12.3.5.3 Body Weight Changes

Body weight changes (gains or losses) for individual animals and groups of animals when compared to concurrent control changes during the course of a study are a criterion of some importance (Heywood 1981; Roubicek et al. 1964; Weil and McCollister 1963). Such changes are usually related to food intake Weight loss may not always be related to toxicity per se (Seefeld and Petersen 1984). The incorporation of the test contaminant into the diet may reduce its palatability.

12.3.5.4 Haematological, Clinical Chemistry, and Urinary Measurements

Regulatory guidelines generally suggest that haematological, clinical chemistry, and urinary parameters be routinely measured in sub-chronic and chronic toxicity studies.

Normal biological variation in inter-animal values and their alteration in response to a variety of inputs means that evaluators will have to contend with much "noise" in this area, and will frequently be presented with scattered, statistically significant effects, in the absence of any evidence of clinically significant relationships to specific toxicity end points. To deal with "noise" there is a need to examine whether the effect noted is within the normal range of variation (concurrent and historical controls). Note that some of these parameters can vary significantly with no clinical manifestations, but others (e.g., serum potassium) have a very narrow normal clinical range and small differences can be important.

Frequently these data show apparently "random" changes in individual group(s) or, less commonly, non dose-related trends in changes across several groups. If using historical control data as an aid to evaluation, only values produced by the identical methods from the same laboratory are valid in such comparisons. Literature values for normal ranges that do not specify the method by which they were obtained should be used with caution.

Sensitivity and specificity of the enzyme changes as diagnostic of organ pathology are greatly influenced by the species selected for testing and are of great importance when specific clinical chemistries are selected for inclusion in toxicity studies. For example, in mammalian species, aspartate transaminase (AST) is not specific to any tissue and thereby elevated plasma AST activity may suggest damage to any one of many tissues. In contrast, alanine transaminase (ALT) is relatively specific to the liver in the cat, dog, ferret, mouse, and rat whereas, in primates, ALT is present in heart, skeletal muscle, and liver.

12.3.5.5 Absolute and Relative Organ Weights

Generally, histopathology is more sensitive for establishing the lowest dose producing an effect than organ or body weight changes. Organ weights are usually reported as absolute organ weights and as relative organ weights (relative to body weight and/or brain weight). Relative organ weight comparisons are used since body weights are often affected by contaminant administration.

A proper evaluation will include consideration of any correlation between organ weights (absolute and relative), histopathological and metabolic/pharmacodynamic data.

12.3.5.6 Post Mortem Observation

Although much progress has been made in the standardisation of nomenclature, to minimise any difficulties in this area, an experienced pathologist will describe each significant lesion type, at least once, in such detail that another competent pathologist can perceive a mental picture of the lesion and form a judgment as to its relevance to the histopathology induced by the contaminant being tested.

12.3.5.7 Analysis and Evaluation of Study Parameters in Toxicity Studies

In this section an analysis and evaluation of study parameters in acute, developmental, reproductive and special toxicity studies is given

Acute Toxicity Studies

Important end-points in acute toxicity studies are clinical signs, gross necropsy signs, and mortality. Since the purpose of acute toxicity studies has moved away from the establishment of a strict, quantitative number for the median lethal dose to an estimate of the likely toxicity range, the emphasis is more on clinical signs and gross organ pathology than on mortality.

Reproductive Toxicity Studies

The critical end-points relate to potential toxic effects on reproductive parameters, including effects on mating behaviour (both sexes), on fertility (both sexes), the implantation of blastocysts, embryonic and fetal development and survival, parturition, lactation, and postnatal survival and development. Thus, a range of reproductive parameters need to be assessed in one or more generations, depending on whether the study is a one-generation (OECD Test Guideline 415), two-generation (OECD Test Guideline 416) or three-generation test. Important end-points to assess within each generation include: time after pairing to mating; mating behaviour; percentage of females pregnant; number of pregnancies going to full term; litter size; number of live births; number of stillborns; pup viability and weight at parturition, and postnatal days 4, 7, 14 and 21 days of age; the fertility index (% of matings resulting in pregnancy); gestation index (% of pregnancies resulting in live litters); viability index (% of pups that survive 4 or more days); and lactation index (% of pups alive at 4 days that survived to day 21, i.e., weaning); gross necropsy and histopathology on some parents (sires and dams), with attention paid to the reproductive organs; and gross necropsy on weanlings

Developmental Toxicity Studies

The critical end-points in developmental toxicity studies relate to potential developmental effects in utero, including death, malformations, functional deficits and developmental delays in fetuses. The following parameters are some of those assessed: number of live litters; number of live fetuses/litter (total and by sex); sex ratio of fetuses; fetal body weights; and litter weights. In addition to the above developmental parameters other reproductive parameters are assessed including the following; number of females pregnant; number of corpora lutea/dam; number of implants/dam; and number and percentage of pre-implantation loss/litter. OECD Test Guideline 414 outlines the protocol for a standard developmental or “teratology” study.

Special Studies

Different classes of contaminants may require special toxicology studies that are not part of the “standard” package of studies. For example, it is common to test organophosphate (OP) pesticides for their ability to cause delayed neuropathy by conducting tests in hens (OECD Test Guideline 419), since this species is especially sensitive to OP pesticide inhibition of neuropathy target esterase.

Toxicokinetic and Metabolism Data

Toxicokinetic (absorption, distribution and elimination) and metabolic data on the handling of the contaminant in the test species, can be very useful in the evaluation and interpretation of sub-chronic and chronic exposure study data.

A number of toxicology text books include chapters on pharmacokinetics and toxicology assessment. The publication, Science and Judgement in Risk Assessment (National Academy of Sciences (NAS)/National Research Council (NRC) 1994), has useful sections on the impact of pharmacokinetic information in Risk Assessment.

12.3.5.8 Interspecies Scaling of Doses

> Where animal bioassays are the source of data, an estimate or measure of the human equivalent dose is required for assessing the health risks posed by contaminant. To derive a human equivalent dose from animal data, the preferred option is to use toxicokinetic data that provides biologically equivalent doses (NHMRC 1999, p. 65).

There are several methods for undertaking scaling of doses including physiologically-based, pharmacokinetic modelling and other models, scaling on a body weight basis, and allometric scaling (body weight raised to the three-quarter power) (ibid). As there is often inadequate data (probably the majority of cases), interspecies differences in physiological and biochemical processes are accounted for by using a generic interspecies factor (ibid).

12.3.5.9 Route-to-Route Scaling

Often the toxicological data are not available for the most appropriate route of exposure for humans. For example, only oral carcinogenicity data may be available, whereas exposure to contaminants by oral, dermal and inhalational routes may be important. Thus, extrapolation from one route of exposure to another may be necessary; this needs to be assessed on a case-by-case basis depending on the available data.

One important consideration in route-to-route extrapolation is determining whether the adverse human health effects are localised to the site of exposure in the body (e.g. carcinomas occurring in the nose or lungs from inhalation) or whether they are a consequence of systemic distribution (e.g. skin cancers arising from arsenic ingestion or some types of leukaemias from benzene inhalation). If the effects are localised to the site of exposure in the body and not a consequence of the systemic distribution of the contaminant, then it may be inappropriate to extrapolate the dose to a different route of exposure. If the effects are consequent to absorption and systemic distribution of the contaminant, then dose scaling between routes of exposure needs to account for the bioavailability of the contaminant by the different routes.

Therefore, bioavailability is an important consideration when extrapolating the applied dose to different routes of exposure. However, additional factors may need to be considered, such as physiological differences between species when extrapolating, for example, from inhalational exposure in animals to oral exposure in humans or vice versa. The assessor should include information about the bioavailability of the contaminant in the experimental studies in the final report.

In cases where bioavailability data are not available, important clues may be gained from the physical and chemical properties and physical state of the contaminant (e.g., liquid, solid or gas) (NHMRC 1999, p. 65).

12.3.5.10 Other Factors in Scaling of Doses

> For inhalational exposure, doses expressed as mg/m^3 or ppm must be converted to mg/kg$_{\mathrm{bw}}$ in the test species by calculations based on the physical properties of the contaminant and minute volumes and respiration rates of the animal (NHMRC 1999, p. 66).

12.3.5.11 Extrapolating Occupational Data to the General Public

Occupational data is often derived from a relatively homogeneous group: usually male, aged between 20 and 65 years and relatively healthy. When applying this data to the general population the differences between the exposed populations should be taken into account as the general population will contain females, and people who are not in the workforce because of their age (young or old) or poor health.

12.3.5.12 Statistical Tests

The objective of a toxicology study is to demonstrate responses of biological importance. Where statistical analyses are used in the assessment process, an awareness of the validity of the test and the degree of certainty (confidence) within the context of the study should be demonstrated.

There are limitations associated with the use of statistics in toxicology (Gad and Weil 1986):

- statistics cannot make poor data better;
- statistical significance may not imply biological significance;
- an effect that may have biological significance may not be statistically significant;
- the lack of statistical significance does not prove safety.

The importance and relevance of any effect observed in a study must be assessed within the limitations imposed by the study design and the species being studied (see also Section 12.2; “Hazard Assessment Part 2: Hazard Identification – Epidemiology”).

If statistical tests have not been used, if inappropriate tests appear to have been used, or if tests not commonly employed have been used, then this should be noted and action taken, e.g., data re-analysis.

A number of textbooks and papers on the application of statistics in experimental toxicology and the life sciences are available; these include Dickens and Robinson (1996), Gad and Weil (1986), Gad and Weil (1989), Lee (1993), Tallarida and Murray (1987) and Waner (1992).

Findings should be considered on the basis of both statistical significance and likely biological significance. The variability of biological data must be remembered in assessing a result as statistically significant. Conversely, a finding that is not statistically significant may have biological significance when considered in the light of the likely toxicological or pharmacological action of the contaminant, or when combined with results from other studies. Thus, evaluators should note trends or transient changes in parameters if there is an indication that these may be related to dosing with the contaminant in some way. This information may be useful when comparing results across studies and in the consideration of the overall significance or relevance of an observed effect, i.e., in one study an effect may be only a trend whilst in another study it may be very clearly treatment-related.

A difficult problem for evaluators is the fact that some studies producing either clearly positive or negative results may have to be considered as flawed. In any long-term study there may be questionable components of the study and an experienced toxicologist must learn to recognise what is useful and reliable and discard what is not. The use of a seriously flawed negative study may provide only a false sense of security. On the other hand, a flawed positive study may be entitled to some weight; how much is a matter of judgement (Task Force of Past Presidents 1982). Data obtained from studies carried out many years ago should not be dismissed out-of-hand simply because they do not meet today's standards; they may provide some useful information. This is a matter for scientific interpretation and judgement on a case-by-case basis.

12.3.5.13 Completion of Hazard Analysis

At this point the assessor should have formulated judgments and supporting rationale concerning:

- the acceptability of the study and its database;
- the existence of biologically important adverse effects;
- the relevance of any factors noted during the evaluation which might have had some bearing on the outcome of the study and modified the findings in some way;
- the likelihood that any of the observed effects were induced by the administered contaminant.

The evaluator should succinctly summarise the critical toxicokinetic and toxicological data, together with any modifying factors for the study under review. The lowest, or most appropriate NOEL/NOAEL, or the absence thereof, should be stated, with a clear indication of the effect(s) on which it was based (i.e., the lowest-observed effect level or LOEL should be apparent). It is important to correlate findings seen in different studies; whilst this is done within the final summary of all toxicity studies, it will often be appropriate to make some mention of cross-study correlations (or the unexpected/unexplained absence of them) within individual study summaries. Possible or proven mechanisms of toxicity should also be discussed and included in the summary.

12.3.6 Evaluation of the Weight-of-Evidence and Consideration of the Toxicology Database In Toto

The essential purpose of toxicity studies is the detection of valid biological evidence of the hazard potential of the contaminant being investigated. In this chapter, the evaluation of the weight of evidence[1] produced by toxicity studies is that process which considers the cumulative data pertinent to arriving at a level of concern about the potential adverse effects of a contaminant. It is composed of a series of judgments concerning the adequacy, validity, and appropriateness of the methods used to produce the data base, and those judgments which bring into causal, complementary, parallel, or reciprocal relationships, all the data considered. Because our knowledge about mechanisms of toxicity is still developing, because good epidemiological evidence is seldom available, and because animal studies are not always conclusive, the information available at a given time may provide only "persuasive" rather than "hard" evidence of a defensible presumption, one way or the other, about the potential health effects of a contaminant under given conditions of exposure. Therefore, it is necessary to succinctly discuss the rationale for judgments and conclusions contained in Risk Assessments together with any associated uncertainties. This becomes important when new data or new scientific knowledge requires re-evaluation of the database or a change in a previous Risk Assessment or regulatory action.

At present, there is no acceptable substitute for informed judgment, based on sound scientific principles, in the analysis, evaluation, interpretation, and weighting of biological and toxicological data derived from animal toxicity studies conducted according to currently available protocols.

It is also accepted practice to apply safety or uncertainty factors to the NOEL/NOAEL derived from animal studies when estimating an ADI (or TDI) as an aid in evaluating the acceptability of actual or potential human exposures

In addition to identifying toxic effects and the doses at which these effects do or do not occur, toxicity studies may yield insight into the mode- or mechanism of action of a contaminant. The evaluator may be able to combine information from a number of studies within the database (e.g., metabolic/toxicokinetic, acute, short-term repeat-dose, subchronic, chronic/carcinogenicity, developmental, reproductive, and genotoxicity studies), to adduce information about the mode or mechanism of toxic action of the contaminant.

At the point of overviewing the entire toxicology database the WHO/IPCS Conceptual Framework for Cancer Risk Assessment (Dybing 2002, Boobis et al. 2006) can be applied. This "Framework" is an analytical tool providing a logical, structured approach to the assessment of the overall weight of evidence for a

---

[1] "Strength of evidence" is commonly taken to mean the degree of conviction regarding the outcome of an experiment e.g. The US National Toxicology Program's "clear evidence", "some evidence", "equivocal evidence" and "no evidence" of carcinogenicity. "Weight of evidence" involves integration of all available data, not just one study.

postulated mode of carcinogenic action. Use of the Framework should increase the transparency of the analysis by ensuring that the facts and reasoning have been documented clearly, including any inconsistencies and uncertainties in the available data.

The Conceptual Framework has been developed to assist in the assessment of carcinogenic end-points, but the principles upon which it is based are broad, and enable its use in analysing modes of action of non-neoplastic effects of contaminants.

12.3.7 Methods for the Hazard Identification of Carcinogens

A variety of Risk Assessment methods has been used elsewhere, for example by the United States Environmental Protection Agency (US EPA 1986, 2005), and the World Health Organization (WHO 1993).

Advances in biological knowledge are enabling mechanistic data, pharmacokinetic data and other relevant data to be increasingly taken into account in classifying and assessing the risks of carcinogens.

Existing methodologies have difficulties in conveying the broad range of health implications of exposure to contaminants. This, combined with a high “dread factor” for cancer, has resulted in many cases in a disproportionate regulatory, political and public focus on cancer as compared to other-than-cancer health effects.

Australia, for example, uses a variety of methods for classifying carcinogens including the International Agency for Research on Cancer’s method for the classification of carcinogens (IARC 1978) which was the first system for qualitatively categorising carcinogenic contaminants Initially, the approach was to adopt a strength-of-evidence scheme to decide whether, for humans and experimental animals separately, there was sufficient or limited evidence of carcinogenicity for a contaminant, mixture, or exposure circumstance, or whether data were inadequate for classification (prior IARC monographs essentially only summarised existing tumourigenicity studies). Since then, the scheme has evolved whereby now all data, including human, animal and in vitro studies are assessed for an overall weight-of-evidence evaluation of human carcinogenicity (Vainio and Wilbourn 1992).

A major contributor to this evolution was the decision that “in the absence of adequate data on humans, it is reasonable, for practical purposes [it is biologically plausible and prudent (IARC 1987)], to regard contaminants for which there is sufficient evidence of carcinogenicity in animals as if they presented a carcinogenic risk to humans” (IARC 1983, 1987). Thus considerable weight is given to the animal cancer bioassays, though some researchers are not convinced of the validity of this philosophy.

A further decision by IARC was to incorporate information on the mechanism of action of contaminants in the evaluation process (Vainio et al. 1992). For example, in practical terms, this means that category Group 1 (sufficient evidence for carcinogenicity in humans) “could be extended to include agents for which the evidence of carcinogenicity in humans is less than sufficient, but for which there is sufficient evidence of carcinogenicity in experimental animals and strong evidence in exposed humans that the contaminant acts through a relevant mechanism of carcinogenesis" (Vainio et al. 1992). This aspect of the evaluation process will become increasingly important as the understanding of mechanistic pathways improves; great advances are being made, especially with the advent of sophisticated laboratory molecular techniques. Essentially four descriptive dimensions of mechanistic data are proposed:

- evidence of genotoxicity (i.e., structural change at the level of the gene);
- evidence of effects on the expression of relevant genes (i.e., functional changes at the intracellular level);
- evidence of relevant effects on cell behavior;
- evidence of time and dose relationships of carcinogenic effects and interactions between contaminant (Fitzgerald 1993, p. 51).

12.3.8 The Hazard Identification Report

The Hazard Assessment component is likely to be based on a number of studies, conducted in different species within each toxicology study type e.g., acute, chronic, developmental, or reproductive toxicity. The toxicity studies [or review(s)/monograph(s)] on which the Hazard Identification and assessment are based should be clearly identified. The report must be transparent, accountable and defensible

12.4 Hazard Identification-Epidemiology

12.4.1 Introduction

Epidemiology and toxicology are complementary in Risk Assessment. Epidemiology is the direct human evidence component and, if based on sound epidemiological methods, can provide the most important evidence in characterising risk.

Epidemiology is "the study of the occurrence and distribution and determinants of health related states or events in specified populations, including the study of the determinants influencing such states and the application of the study to the control of health problems" (Porta 2008).

An excellent introductory text is:

Bonita R, Beaglehole, Kjellstrom T (2006). Basic epidemiology. 2nd edition. World Health Organization, Geneva.

Epidemiological methods are used to investigate the cause of adverse human health effects; the natural history of health conditions; the description of the health status of populations; and to evaluate health related interventions (Beaglehole et al. 1993). In the context of Human Health Risk Assessment, epidemiological methods may also be used to characterise population exposures, investigate perceived clusters of disease, to develop health surveillance programs to establish a baseline, and to monitor the consequences of Risk Management activities.

Epidemiology can assist Hazard Assessment in both Hazard Identification and Dose-response assessment.

There are often unrealistic expectations of what an epidemiological study may be able to achieve. As Mundt et al. (1998) noted, if the limitations of epidemiological studies are not understood by the Risk Assessment team, the validity of an assessment might be compromised by including inappropriate, possibly misleading, epidemiological data. The systematic appraisal of epidemiological studies is intended to answer the question "Is there any other way of explaining the set of facts before us [i.e., the study results], is there any other answer equally, or more, likely than cause and effect?" (Hill 1965 in WHO 2000). Alternative explanations may result from chance, bias and confounding (WHO 2000).

12.4.2 Bias and Confounding: Key Concepts in Environmental Epidemiology

There are many ways in which error can be introduced into epidemiological studies. Error may be random (due to chance alone, and potentially reduced by improving sample size), or systematic (and not reduced by increasing sample size). There are two key concepts of systematic error, bias and confounding. The size of the statistical confidence intervals will provide an indication of the potential for random sampling error, but statistical confidence intervals do not represent uncertainty arising from bias or confounding.

Bias occurs if there is a systematic tendency by a study to produce results that diverge from the truth. There are many sources and varieties of bias, but the most important include selection bias and measurement (or classification) bias. The reader is referred to Bonita et al. (2006) for a succinct account of bias. It may be difficult to precisely estimate the effect bias has in a study, but it is vital for risk assessors to look for and attempt to identify the potential size and direction of bias in interpreting a study's findings.

Confounding is the distortion of the effect of the contaminant of interest by an extraneous factor (Moolgavkar et al. 1999). This may occur if another exposure exists in the study population that is associated with both the disease (or outcome) and the exposure being studied, e.g., a third factor ("confounding variable") that independently affects the risk of developing the disease.

There are specific approaches for the control of confounding that can be used in both the design and analysis of analytic studies providing that the confounding variables have been identified and measured.

12.4.3 Types of Epidemiological Study – An Overview

Broadly speaking, epidemiological activity can be either “descriptive” (reporting and describing the distribution of exposure and effect) or “analytical” (designed to analyse and understand the degree of association between exposure and effect). Descriptive studies include case reports, case series and cross-sectional surveys. Cross-sectional surveys measure exposure and effect in an individual at the same point in time and thus are unable to support causal inference.

In practical terms in environmental epidemiology there are four main categories of analytical study (Moolgavkar et al. 1999):

- cohort (longitudinal) studies;
- case-control studies;
- cross-sectional studies;
- ecological studies (including a special subgroup known as time-series studies).

Cohort, cross-sectional and case control studies differ from ecological studies in that information on exposure and disease is available on an individual basis. With ecological studies this information is only available on a group basis, so the community or region is the unit of analysis.

In case-control studies, exposure and other attributes of cases of the disease under investigation are compared with those from a suitable control or comparison group of persons unaffected by the disease, and analysed to yield effect estimates. The selection of appropriate controls to avoid bias is a significant challenge with case-control studies. They are relatively inexpensive, ideal for studying rare diseases and useful for investigating multiple, different exposures (Gregg 1996).

Cross-sectional studies measure prevalence of disease and measure exposure and effect at the same time. They are relatively easy and economical to conduct and are particularly useful for measuring fixed characteristics of individuals such as socioeconomic status (Beaglehole et al. 1993).

Cohort studies follow cohorts or groups of individuals, defined in terms of their exposures, over time to see if there are differences in the development of new cases of the disease of interest (or other health outcome) between the groups with and without exposure. Such studies can be carried out by either reviewing past records (retrospective) or by tracking people into the future (prospective cohort). The essential feature of these longitudinal studies is that for each individual prior exposure information can be related to subsequent disease experience (Breslow and Day 1987).

Ecological studies involve the investigation of a group of people such as those living within a geographical area such as a region or state. For example, place and time of residence may be used to create surrogate measures of the real exposure of interest (Elliott et al. 1992). Rates of disease and average exposure levels to a particular contaminant are determined independently, and on a group basis. This may give rise to spurious apparent correlation, called the ecological fallacy. Because it is not ascertained whether individuals who have been exposed to the contaminant are the same individuals who developed the disease, statements about causal associations are inappropriate. However, ecological studies are relatively inexpensive for linking available health data sets and environmental information and are useful for hypothesis-generation (Yassi et al. 2001). Examples of ecological studies are the assessments of the relationship between tobacco sales in different countries and lung cancer rates, and fluoride in water supplies and dental caries.

A subset of ecological studies, known as time series studies, is regarded as very helpful in understanding the influence of short-term fluctuations in air contaminants on day-to-day changes in population morbidity and mortality after controlling for factors such as season and air temperature. However, disentangling the effects of individual contaminants as measured in a mixture such as urban air pollution can be quite difficult.

To strengthen the design of ecological studies, Nurminen (1995) recommended the selection of areas with populations that:

- are homogeneously exposed (to minimise within-area exposure variation);
- represent different extremes of exposure distribution (to maximise between-area exposure variations);
- are comparable with respect to co-variate distributions (e.g., socio-economic status, demography); and
- use the smallest possible sampling units for ecological analysis.

The largest number of environmental epidemiology studies found in the literature is of the ecological or cross-sectional type, because they are easier to carry out and cost less (Thomas and Hrudey 1997). However, as noted above and discussed further below in relation to assessment of causality, such studies may be useful for identifying potential hazards or hypothesis generation, but they cannot determine cause and effect.

Characteristics of the various study types are summarised in Table 12.1. Epidemiological studies are rarely definitive and a single epidemiological study cannot establish causality. A "weight of evidence" approach is generally required, involving the interpretation of integrated information.

Unfortunately experimental interventions such as randomised controlled trials are rarely available to assist Human Health Risk Assessment.

Table 12.1 Advantages and disadvantages of different observational study designs

	Ecological	Cross-sectional	Case-control	Cohort
Probability of:				
selection bias	N/A	Medium	High	Low
recall bias	N/A	High	High	Low
loss to follow-up	N/A	N/A	Low	High
confounding	High	Medium	Medium	Low
Time required	Low	Medium	Medium	High
Cost	Low	Medium	Medium	High

Beaglehole et al. (1993, p. 42)

Epidemiological studies, depending on their design, may serve two purposes; hypothesis-generation or assessment of a causal relationship. Their ability to evaluate a causal relationship may be limited by a lack of control of potential confounders or a lack of power (which is usually the result of limited sample sizes) (Samet et al. 1998).

12.4.3.1 Observational Studies

Different observational study designs have different applications, advantages and disadvantages (see Table 12.2). These comparisons assume the different types of studies are equally well designed. Even so, design variations may affect their performance and provide exceptions. See Beaglehole et al. (1993) for a more detailed description.

Table 12.2 Guidelines for the assessment of causation

Temporal relation	Does the cause precede the effect (essential)
Plausibility	Is the association consistent with other knowledge? (mechanism of action; evidence from experimental animals)
Consistency	Have similar results been shown in other studies?
Strength	What is the strength of the association between the cause and the effect? (In general, relative risks greater than 2 can be considered strong)
Dose-response relationship	Is increased exposure to the possible cause associated with increased effect?
Reversibility	Does the removal of a possible cause lead to reduction of disease risk?
Study design	Is the evidence based on a strong study design?
Judging the evidence	How many lines of evidence lead to the conclusion?

Beaglehole et al. (1993, p. 76)

12.4.4 Assessing the Relationship Between a Possible Cause and an Outcome

A cause is “an event, condition, characteristic or a combination of these factors which plays an important role in producing the disease” (Beaglehole et al. 1993, p. 76).

Causation of adverse health effects is affected by four types of factor:

- *predisposing factors* such as immune deficiencies, gender and previous illness;
- *enabling factors* such as poor nutrition and bad housing may favour the development of disease;
- *precipitating factors* such as the exposure to a specific disease contaminant; and
- *reinforcing factors* such as repeated exposure which may aggravate an established disease or state (Beaglehole et al. 1993).

The term “risk factor” is commonly used to describe factors that are positively associated with the risk of development of a disease, but that are not sufficient in themselves to cause the disease. A “sufficient” cause is one which “inevitably produces or initiates a disease” and a “necessary” cause is one for which “a disease cannot develop in its absence” (Beaglehole et al. 1993). In the biological sciences there is often a constellation of components acting in concert for a cause to create an effect, and many of the components of a “sufficient cause” may be unknown (Rothman and Greenland 1997). At the low levels of exposure commonly encountered in the environment and where there may be a range of contributory factors present, it may be difficult or inappropriate to assign this nomenclature to a contaminant even though the contaminant is accepted as causing a specific effect with high exposures.

As with other scientific disciplines, epidemiology has attempted to define a set of causal criteria to help distinguish causal from non-causal associations. In the first place other explanations for a potentially causal association must be excluded (such as chance, selection or measurement bias, or confounding. Particularly rigorous scrutiny should be given to studies giving a positive but not statistically significant result. Figure 12.2 illustrates this process.

If alternative explanations such as bias and confounding can be excluded, it is then useful to systematically apply guidelines for assessing causation from Beaglehole et al. (1993) as shown in Table 12.2. The concepts in these guidelines derive from work by Hill (1965) and others. However, as Rothman and Greenland

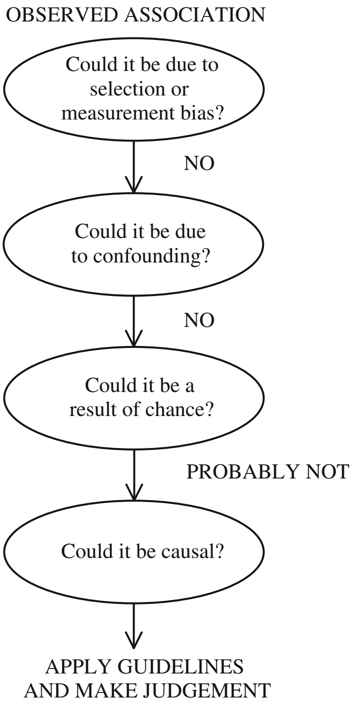


Fig. 12.2 Assessing the relationship between a possible cause and an outcome when an association is observed (Beaglehole et al. 1993, p. 75)

(1997) note, apart from temporality (whereby a putative cause must precede the effect) there are no necessary and sufficient criteria for determining whether an observed association is causal. Thus the term "guidelines" is more appropriate than the slightly more absolute "criteria"; and there is not necessarily an easy epidemiological road-map to finally determine causation.

With environmental health in particular, much decision-making rests on a "weight of evidence" approach rather than definitive proof of cause, which is commonly not available – hence the final concept, "Judging the evidence" in Table 12.2, is particularly relevant.

These guidelines are ordered in a logical sequence for making judgements on causality. They are not weighted equally, and their relative contribution to a final judgement will vary from one situation to another (Thomas and Hrudey 1997).

Consistency can be demonstrated if several studies give the same result, especially if a variety of designs is used in different settings since this reduces the likelihood that all studies are making the same mistake. However, other factors such as different exposure levels or study conditions may need to be taken into account, and the best-designed studies should be given the greatest weight. It is important to note that in environmental epidemiology, reliance on a single pivotal study is the exception rather than the rule.

The technique of meta-analysis grew out of the need to reduce random error in clinical trials. Meta-analysis in the context of systematic reviews can be used to pool the data from well-designed studies, each of which may deal with a relatively small sample size, in order to obtain a better overall estimate of effect. Meta-analysis has pitfalls if poor quality studies are included, and needs to be applied with caution to observational studies – which are less able to control for confounding than randomised trials. Standard methods for conducting and reporting systematic reviews have been published (Greenhalgh 1997). The reader is also referred to an excellent resource published by NHMRC (2000), "How to review the evidence: systematic identification and review of the scientific literature".

The strongest evidence comes from well-designed and competently conducted randomised controlled trials. The National Health and Medical Research Council (1999) places strongest emphasis on evidence obtained from systematic reviews of all relevant (and well-conducted) randomised controlled trials ("level 1").[2]

However, there are relatively few such trials available for environmental health hazards that could form the basis for a systematic review. Most apply to the effects of treatment or prevention campaigns. A rare example is the Melbourne Water Quality Study which was a blinded study involving real and sham domestic reverse osmosis water filters and an assessment of acute gastrointestinal disease (Hellard 1999).

In practice, most evidence comes from observational studies (e.g., nearly all the evidence on the health effects of smoking). In well-conducted cohort studies bias is

---

[2] The NHMRC document is oriented towards clinical interventions and clinical practice guideline development. More recently, advice to guide assessment of epidemiological evidence for environmental health practice has been provided in NHMRC (2006). Ambient Air Quality Standards Setting: An Approach to Health-Based Hazard Assessment.

minimised. Case control studies are subject to several forms of bias and weaknesses related to time-sequence but, if well designed, may still provide useful evidence for the causal nature of an association. Cross-sectional studies are weaker as they provide no direct evidence on the time sequence of events.

Ecological studies are the least satisfactory because of the dangers of incorrect extrapolation to individuals from data derived from regions or countries. However, where certain exposures cannot normally be measured individually (e.g., air pollution, pesticides residues in food, fluoride in drinking water) evidence from ecological studies may be important in environmental health decision making (Beaglehole et al. 1993). Time-series studies demonstrating health outcomes associated with fluctuating air contaminant levels may be one particularly useful example.

The ranking in this Table 12.3 assumes that studies are well designed and well conducted in each case. Even the presence of "a strong ability to 'prove' causation" should be supplemented by mechanistic knowledge to be confident of causation.

Table 12.3 Relative ability of different types of study to "prove" causation

Type of study	Ability to "prove" causation
Randomised controlled trials	Strong
Cohort studies	Moderate
Case-control studies	Weak/Moderate
Cross-sectional studies	Weak
Ecological studies	Weak

Adapted from Beaglehole et al. (1993)

12.4.5 The Strengths and Limitations of Observational Epidemiology Versus Experimental Toxicology

Epidemiological studies are crucial for assessing effects directly in humans and estimating population attributable risks. However, their power of resolution is limited, mainly because of the difficulties in estimating exposure precisely and in controlling bias. Toxicological studies are necessary for elucidating causal mechanisms, which may be important for determining dose-response relations and extrapolating to low doses in Risk Assessment, but direct generalisations to humans based on animal data are often uncertain (Pershagen 1999).

Epidemiological studies are often given increased weighting because they come from humans but, compared to toxicological studies of animals, may be more costly and time consuming and more likely to result in ambiguous findings (Samet et al. 1998). However, substantive findings have been obtained at times through opportunistic studies of highly exposed groups – such as occupational cohorts or communities that have been inadvertently exposed to contaminants e.g., via food or water. These can be either observational epidemiological studies, or what Lilienfield and Lilienfield (1980) called “natural experiments”.

12.4.5.1 Hazard Identification

Epidemiology has a number of potential advantages over animal toxicology in the area of Hazard Identification:

- it directly assesses human health risk;
- absorption, metabolism, detoxification and excretion may vary between the human species and the animal species studied and does not need to be taken into account in epidemiological studies;
- sample sizes for human studies may be much larger than those feasible for animal studies;
- genetic diversity may be greater in humans compared to the selected animal strains used in toxicological studies;
- epidemiological studies may include different groups (e.g., the young, old and susceptible) that may not be included in the usually relatively homogeneous groups used in toxicological studies;
- effects on some aspects of mental function or behaviour, and more subjective effects such as nausea or headache, can be better assessed in human studies.

Roseman (1998) and Samet et al. (1998) provide further advantages:

- reduced uncertainty about interspecies variability in metabolism, lifespan, and genetic diversity;
- complex temporal patterns of exposure and doses in situations requiring Risk Assessment may be impossible to replicate in animal studies; whereas some epidemiological studies may be more useful for understanding these complex dose-response relationships;
- the ability to assess large numbers of people exposed to low levels of a contaminant. The doses from exposure to a hazardous contaminant in epidemiological studies are often considerably less than in toxicological studies. This may have the advantage of providing information about the exposure range of interest although, if they are the result of (prolonged) adult occupational exposures, the exposures are likely to be considerably more than those experienced by people in the general population. With appropriate tools small differences in relative risk in large populations may be able to be assessed.

However, epidemiological studies are often limited by the amount and quality of data available on dose and tend to address exposure-response relationships (i.e., they are based on whether or not exposure occurred) rather than dose-response relationships. Saunders et al. (1997) reviewed 14 key relevant studies selected from a short list of 43 analytical studies assessing human health effects in relation to hazardous waste sites, and found that poor exposure measurement was a major factor in the overall lack of convincing evidence of causation from these studies. It is often the case that only a broad indication of the level or nature of exposure may be deduced from epidemiological studies.

Quantitative description of dose-response relationships may be hampered by incomplete information on exposure (especially for biologically relevant time windows), by exposure or dose misclassification, or by the use of surrogate markers of exposure. Incorrect information about the exposure may bias the description of the exposure-response relationship. If there are wide confidence intervals around the results there can be substantially different policy endpoints depending on whether the upper bound, the lower bound or the midpoint has been chosen for policy making (Samet et al. 1998).

Commonly too there are insufficient epidemiological data to discriminate between alternative models that could describe the dose-response relationship. This is particularly important at very low exposure levels and this is where both epidemiological and toxicological data are often limited.

The reviewer or risk assessor should answer the basic question of whether the epidemiologic data, in an individual study or cumulatively, are adequate for use in dose-response evaluation. There is no formula or quantitative weighting scheme prescribed for making this judgement.

If epidemiologic data adequate for dose-response evaluation are not available, and a Risk Assessment is being developed for use in making an important regulatory decision, and if it is feasible to develop new epidemiologic data, or to extract new data from existing studies, an effort should be made to develop and provide good epidemiologic dose-response data that can be used together with, or in preference to, high-dose animal data.

12.4.6 Undertaking Health Studies

In some situations there will be a need to undertake health studies as part of a Risk Assessment. The design of health studies should be underpinned by epidemiological principles. A range of factors needs to be considered before embarking on a health study (ATSDR (1996):

- public health significance including the scientific value for the Risk Assessment and generalisation to other situations;
- community perspective and involvement;
- ability to provide definitive results;
- resource availability;
- support from local, state and federal agencies

The first level of health studies explore or generate hypotheses about exposure-outcome associations and address specific exposures, community health concerns, or specific information needs. Examples are:

- cross-sectional studies;
- pilot;

- cluster investigations;
- comprehensive case reviews;
- situation-specific surveillance;
- health statistics reviews;
- exposure investigations;
- disease and symptom prevalence surveys.

The quality of these studies depends on (ATSDR 1996):

- a reasonable ability to document and characterise exposure in the target area;
- an adequate study size for the type of study recommended;
- an ability to identify and locate subjects and records;
- appropriate comparisons for rates of occurrence or levels of exposure; and
- an ability to control confounding factors and biases (when possible).

More complex health studies are specifically designed to test scientific hypotheses about the associations between adverse health outcomes and exposure to contaminants in the environment. Examples are (ATSDR 1996):

- case-control studies;
- cohort;
- nested case-control.

The quality of these studies depends on (ATSDR 1996):

- an ability to reasonably estimate or document individual exposures;
- an ability to document or validate human health outcomes;
- an adequate study size and statistical power;
- an ability to identify and locate subjects and records;
- availability of an appropriate control or comparison population;
- an ability to control confounding factors and minimise biases; and
- an ability to determine influence of environmental, behavioural, or other factors.

12.5 Dose-Response Assessment

12.5.1 Introduction

The following section uses material from the NHMRC's Toxicity Assessment Guidelines for Carcinogenic Soil Contaminants (1999) and Klaassen (1996).

There are different ways of characterising dose response relationships including

- effect levels (e.g., LD_{50}, LC_{50}, ED_{10}) and no observed adverse effect levels (NOAELs);
- margins of safety;

- therapeutic indices; and
- models to interpolate high dose experimental data to the low doses likely to be experienced in the environment (Klaassen 1996).

There are often limited human exposure data and animal bio-assay data are most often used for dose response assessment. The use of these data requires extrapolations from animals to humans and from high doses to low doses (ibid).

Where data is derived from animal studies, the doses are nearly always higher (often considerably higher) than the experimental range. The shape of the dose response curve below the experimental range will be unknown and can have a variety of shapes depending on the mathematical model used for fitting the curve. The choice of the model should be based on mechanistic information about how the contaminant exerts its effects if such information is available.

The dose response curves for different effects will have different shapes and will occur at different doses (see Fig. 12.3). The shape of the dose response curve will be different again when dealing with, for example, an essential trace element such as copper where at low doses there will be a dose response curve for the effects of deficiency and at higher doses another dose response curve describing the effects of excess.

The International Programme on Chemical Safety has reviewed many of the issues relating to dose-response in a recent publication (WHO 2009).

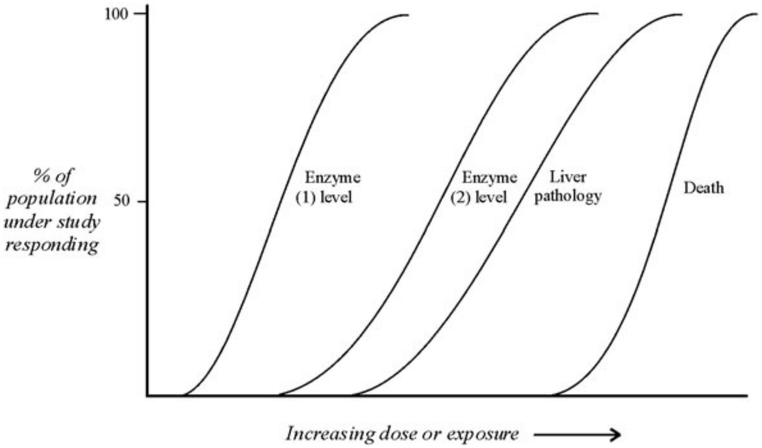


Fig. 12.3 Different dose-response curves for different effects from a hypothetical contaminant

12.5.2 Methodologies

All methodologies make the distinction between neoplastic and non-neoplastic endpoints in Risk Assessment. The impetus for this distinction was the concept of a lack of threshold in the dose-response for carcinogens based on the initial premise that all carcinogens are mutagens (Ames et al. 1973). One mutation or one DNA damage event was considered sufficient to initiate the process that leads to the development of cancer. In contrast, the dose-response was assumed to have a threshold for non-cancer effects, the assumption being that, for non-cancer effects, there is a dose below which the risk of adverse effects will be nil; in effect, that there is a "safe" dose.

In more recent times, as it has become evident that not all carcinogens have genotoxicity as their prime mode of action (Ashby and Tennant 1991), the dose-response curve has often been assumed to be non-threshold for genotoxic carcinogens and threshold for non-genotoxic carcinogens in countries such as Canada and some European countries (Whysner and Williams 1992; Health Council of The Netherlands 1994) and in the WHO Drinking Water Guidelines (WHO 1993). However, the WHO considered that these approaches are not suitable to the development of generic guidance values in Environmental Health Criteria documents because they "…require socio-political judgements of acceptable health risks" (WHO 1994).

In these examples, the distinction between a genotoxic carcinogen and a non-genotoxic carcinogen is a science policy decision for regulatory purposes and does not necessarily reflect the mechanism of carcinogenesis. It does not mean that a non-genotoxic carcinogen does not affect the genetic material of the cell under some circumstances, nor that a genotoxic effect is the only event required for the development of cancer by a genotoxic carcinogen.

With advances in biological knowledge, mechanistic data, pharmacokinetic data and other relevant data are increasingly being taken into account in classifying and assessing the risks of carcinogens. The US EPA has revised its guidelines for carcinogen Risk Assessment (US EPA 2005) and they stress the importance of "a critical analysis of all of the available information that is relevant to assessing the carcinogenic risk as the starting point from which a default option may be invoked if needed to address uncertainty or the absence of critical information" (p. 1.11).The use of mode of action in the assessment of potential carcinogens is a main focus of these cancer guidelines as a result of "the significant scientific advances that have developed concerning the causes of cancer induction" (p. 1.10). The guidelines also emphasise uncertainty analysis in the Risk Assessment process and "the importance of weighing all of the evidence in reaching conclusions about the human carcinogenic potential of contaminant" (p. 1.7).

12.5.3 Threshold Approaches

A threshold is considered to occur because of biological mechanisms such as the ability to metabolise or excrete a toxin or to repair damage up to a certain dose.

The approach with these models is to derive exposure limits such as an ADI, a Provisional Tolerable Weekly Intake (PTWI), Tolerable Daily Intake (TDI) or RfD (Barnes and Dourson 1988; Dourson et al. 1996; WHO 1994). This approach makes no attempt to calculate a level of risk at low exposures. Rather, it derives a dose which is apparently without effect in a human population or suitable animal model, and then applies a factor to derive an exposure which has a high likelihood that no effect will occur in the general human population.

These exposure limits are derived by first determining the No Observed Adverse Effect Level (NOAEL) or, if the NOAEL cannot be determined, the Lowest Observed Adverse Effect Level (LOAEL) and dividing the value by factors to account for:

- interspecies differences (extrapolating from animals to humans);
- intraspecies differences (differing sensitivities between individuals);
- the severity of the adverse effect; and
- the quantity and quality of the scientific data.

Traditionally, safety factors for intraspecies and interspecies differences have each been assigned values of ten, and the other two have been assigned values between 1 and 10. An additional factor of ten is sometimes used if the NOAEL was not established in the study. The individual factors are then multiplied to determine an overall safety factor by which the NOAEL is divided to give the ADI, PTWI, TDI or RfD.

Historically, the most common overall factor used by a number of regulatory bodies is 100 if a large toxicological database has been assessed, although the overall factor can range from 10 to 10,000. From the data available on humans and experimental animals, it appears that interspecies and intraspecies differences are in general less than 10, hence the often-used overall safety factor of 100 for these two factors is conservative and adequately protective of public health (Johannsen 1990; Renwick and Walker 1993).

The decision on the magnitude of factors to use is predominantly based on expert or informed judgement. Whilst this approach to selecting the number and magnitude of the safety factors appears to be arbitrary, current knowledge of the biological processes which cause inter– and intraspecies variation (e.g., metabolic and other pharmacokinetic rate differences) support the choice of safety factors.

REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) provides a European Union legislative framework for chemicals. Both REACH and WHO are producing documents on factors derived from empiric databases and substance – specific adjustment factors.

12.5.4 Non-Threshold Approaches

These approaches do not recognise the possibility of a threshold effect and are appropriate for radiation and for some genotoxic carcinogens. It is, as a science policy decision, applied to all carcinogens by the US EPA.

Non-threshold models assume linearity between the lowest experimentally derived dose and the zero dose. This implies that there is a calculable probability of an adverse effect (risk) no matter how small the dose.

Numerical estimates of risk probabilities are generated by fitting one or more mathematical models to the data in the experimental dose range and extrapolating to the low environmental exposure doses. For example, low-dose extrapolation using a linear model is a default approach for cancer Risk Assessment in the USA (US EPA 1996) and is one approach which has been used by the WHO for genotoxic carcinogens in deriving drinking water guidelines (WHO 1993).

12.5.5 Threshold Versus Non-Threshold Approaches

This area of scientific debate largely centres on the management of carcinogens.

The important conceptual distinction between non-threshold methods and those which derive an acceptable exposure from the NOAEL using a safety factor is that the safety factor approach makes no attempt to determine a finite level of risk at low exposures whereas the linear methods make an estimate of the risk at low exposures. The NOAEL is assumed to be the threshold dose for the effect. Both approaches have advantages and disadvantages.

The advantages of the threshold approach are that the NOAEL is relatively easy to determine, and the process is simple to use, easy to understand and allows the use of expert judgement. In the few cases where epidemiological data have become available, the ADIs derived by this method have been validated (Lu and Sielken 1991). Additionally, the approach has been applied seemingly in a consistent fashion by the WHO in the last three decades in deriving ADIs for pesticides (Lu 1995). The safety factor approach remained essentially unchanged until 1994 (WHO 1994), although a number of articles were published suggesting modifications or improvements (e.g., Calabrese and Baldwin 1994; Calabrese and Gilbert 1993; Crump 1984; Johannsen 1990; Lewis et al. 1990; Lu and Sielken 1991; Zbinden 1979).

Because it provides numerical estimates of risk at all doses, the non-threshold approach, in principle, has the potential advantages (if the estimates are correct) of allowing: computation of comparative risks in the sub-experimental range, which may be a useful tool in Risk Management and communication; potency comparisons between contaminant at a particular risk level; and estimates of the increased risks if a particular dose is exceeded. It has been argued (McMichael 1991) that risk estimates by this approach approximate those seen in humans in some cases and where there are disparities they are overestimates of the risks.

Both the threshold and non-threshold methods, however, are likely to be unduly influenced by the selection of doses. The choice of the NOAEL is limited to one of the doses included in the experimental design. The biological no effect dose may occur at this dose or at a dose not included in the study. The closeness with which the selected NOAEL truly reflects the actual no effect dose has an obvious impact on the degree of protectiveness in the derived ADI, PTWI or RfD. Furthermore, the NOAEL is influenced by the biological effects monitored, the number of animals in the test groups, the spontaneous incidence of the adverse effect, and the criteria used to determine when the incidence in a test group exceeds that in the controls (Renwick and Walker 1993).

Additional limitations of the threshold approach include: the NOAEL is often perceived as a biological threshold, whereas it is a threshold limited by the experimental protocol; risk is expressed as a fraction of the guidance dose (e.g., ADI); it makes limited use of the dose-response slope; the choice of safety factors has been arbitrary to some extent and the process does not generate a range of estimates of risk, but rather a single estimate of a dose below which no adverse effects are likely to be produced.

Dose selection in non-threshold models has been discussed by Lovell and Thomas (1996) who suggest that the estimate of q_1^* (the 95% upper confidence limit of the slope estimate used for the linear multi-stage model used by the US EPA) is so dependent on the doses selected that it is almost independent of, or at least insensitive to, the actual tumour incidences in the dose groups. Specifically, the highest dose in an animal bioassay has overwhelming influence on the estimate of q_1^*, thus leading to the overestimation of risk at very low doses, with the extent of overestimation increasing as the environmental exposure becomes lower. Typically, the highest dose in a carcinogenicity bioassay is the maximum tolerated dose (MTD), a dose that causes no more than a ten percent decrease in body weight and no other overt toxicity. The MTD is very much greater than doses expected from non-occupational environmental exposures. Therefore, the dose which is the least relevant to environmental Risk Assessment has the greatest influence on low dose risk estimates.

Non-threshold models currently in use are inflexible and generally do not take account of the complexities of the events between exposure to a contaminant and the induction of a neoplasm. Risks estimated at doses below the range of experimental data can vary considerably depending on the model used, even though the various mathematical models used generally fit the experimental data equally well (Crump 1985; Paustenbach 1995). The numerical expression of the estimated level of risk falsely gives the impression that it represents an exact measure of actual risk. This numerical expression provides little or no information on the uncertainties related to the estimated level of risk, nor does it allow comparison with values for non-cancer health effects.

Low-dose linearity assumes a positive slope of the dose-response curve at zero dose and implies that a single, irreversible genetic event at the initiation stage of carcinogenesis leading to transformation of a cell, is sufficient by itself to lead to the development of cancer. The major difficulty in this debate is the impossibility of testing experimentally the shape of the dose-response curve at extremely low doses (Purchase and Auton 1995).

A transformed cell which has acquired the potential to develop into a tumour, will probably realise that potential only rarely (US EPA 1996), most likely because of the natural large scale repair of DNA damage and other defence mechanisms of the body (DOH 1991). Furthermore, whilst it is generally accepted that mutagens and mutations play a role in the development of cancer, carcinogenesis is more than mutagenesis, with a number of non-mutagenic as well as mutagenic events taking place during the process (Bishop 1991). The shape of the dose-response curve at any one of these steps, not just the mutagenic events, can influence the shape of the dose-response curve for the carcinogenic response. Factors, such as genetic make-up, lifestyle and other environmental factors, may also have a modifying influence on the processes of carcinogenesis.

12.5.6 Mechanistically-Derived Models

These use models, which describe biological mechanisms by mathematical equations. They assume that the toxic effect results from the random occurrence of one or more biological events. These are known as stochastic events (Klaassen 1996). Examples of mechanistically-derived models are the One-hit, Log-probit, Weibull, Linearised Multi-stage and Moolgavkar-Venson-Knudson models.

Mechanistically-derived models have been particularly used for cancer modelling and especially those based on radiation exposures. The simplest form is a "one hit" linear model in which only one "hit" or critical cellular interaction results in the alteration of a cell. This model would propose that a single molecule of a genotoxic carcinogen would have a "minute but finite chance of causing a mutational event" (Klaassen 1996). From these models more complex models based on multihits or multistage events have been derived. Although conceptually based on biological mechanisms, most of these models do not rely on independently validated parameters describing the mechanisms, but rely on fitting curves to empirically observed data.

More recently these models have been adapted to take into account information based on knowledge of the relevant physiology and toxicokinetics (Physiologically-based toxicokinetics (or pharmacokinetics) modelling). These models take into account the effective dose at the target organ. A further development has been to make generalised mechanistic models take into account specific biological processes such as the Moolgavkar-Venson-Knudson model that uses a two-stage model for carcinogenesis (Klaassen 1996).

12.5.7 Benchmark Dose Approach

The benchmark dose (BMD) approach has been used in dealing with both cancer and non-cancer end points. It is described in EHC170 and a modified version for use with carcinogenic soil contaminants is described in "Toxicity Assessment Guidelines for Carcinogenic Soil Contaminants" (NHMRC 1999). The benchmark dose corresponds to a predetermined increase (usually 5%) of a defined effect in a test population. Figure 12.4 illustrates how it is derived using a 5% increase. Mathematically it is the statistical lower confidence limit on the dose that corresponds to that predetermined increase although some agencies are using a best estimate rather than a lower confidence limit (IEH 1999).

In this example, $LED_5 = BD$, and LED_5 is the lower confidence limit of the effective dose causing a 5% increase in a defined effect.

For developmental toxicity the BMD_5 values have been similar to statistically-derived NOAELs for a wide variety of developmental toxicity end points (Klaassen

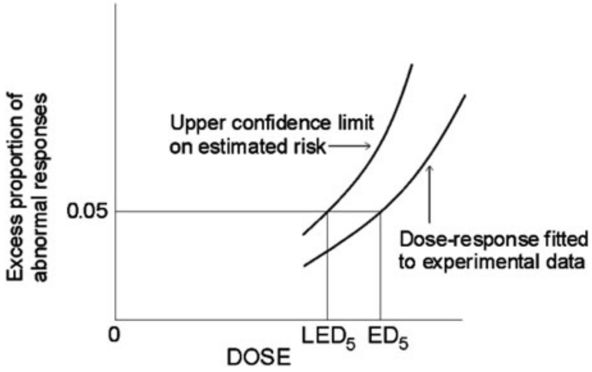


Fig. 12.4 Graphical illustration of the benchmark dose approach. Adapted from WHO (1994)

1996). BMD approaches are also being developed and tested in regard to acute inhalation toxicity (Fowle et al. 1999), to the relationship between the BMD and the MTD (Gaylor and Gold 1998), and to addressing statistical procedures available for calculating BMDs and their confidence limits for non-cancer endpoints (Gaylor et al. 1998).

Particular advantages of the BMD approach include:

- taking into account information from the entire dose response curve rather than focussing on a single test dose such as is done with the NOAEL approach;
- the use of responses within or near the experimental range versus relying on extrapolations to doses considerably below the experimental range;
- the use of a consistent benchmark response level that crosses a range of studies and endpoints;
- it is less influenced than NOAEL approaches by the arbitrary selection of doses (Crump 1984);
- it is able to be rigorously described;
- it uses all available relevant information.

Its disadvantages are that it may not be possible to define the shape of the dose response curve because of limited dose groups or the number of animals per group and it also requires greater statistical expertise than the NOAEL type approach (IEH 1999).

Use of a benchmark dose with 5% extra risk provides a more data sensitive and less model sensitive endpoint than using 1% extra risk (Klaassen 1996, NHMRC 1999).

When the benchmark response is within or near the experimental range of data the corresponding values of the benchmark doses are not greatly sensitive to the choice of the model used, but the best scientific choice of a model would be a biologically-based mechanistic model.

12.5.8 Inter- and Intra-Species Considerations

The material in this section is from WHO (1999, p. 20) with slight adaptation.

The strains and species of laboratory animals exposed in toxicity studies have been selected to show minimum inter-individual variability. Compared to laboratory animals, humans represent a very heterogeneous population with both genetic and acquired diversity.

Sources of inter-species and inter-individual variations in toxicokinetics include: differences in anatomy (e.g., gastrointestinal structure and function); physiological function (e.g., cardiac output, renal and hepatic blood, glomerular filtration rate and gastric pH); and biochemical differences in, for example, enzymes involved in xenobiotic metabolism.

In some cases, it may be possible to conclude that effects detected in animals are unlikely to be relevant to humans. In other cases, there may be data to indicate that humans are likely to be more or less sensitive than animal species; this information is important for consideration in the selection of critical effects.

In extrapolating between species, three aspects need to be considered: (1) differences in body size, which require dose normalisation or scaling (often done by expressing the dose per kg body weight); (2) differences in toxicokinetics, particularly bioactivation and/or detoxification processes; and (3) the nature and severity of the target for toxicity.

The greater potential variability in heterogeneous human populations must be addressed in Risk Assessment. Sources of inter-individual variability in human populations include, for example, variations in genetic composition, nutrition, disease state and lifestyle.

12.5.9 Mixtures

While important, currently there is no agreed international approach to assessing mixtures of contaminant. Where data (including mechanistic data) are available on the interaction of contaminant this should be taken into account in the Risk Assessment.

The Agency of Toxic Substances and Disease Registry (ATSDR) in the US uses one approach that includes performing a critical synthesis of relevant data and then identifying generalisable rules that can be used in site-specific assessments of health risk following exposure to mixtures. This approach allows research to: identify what contaminant mixtures may affect public health; evaluate the potential for exposure of human populations to contaminant mixtures; study the pharmacokinetic behaviour of contaminant mixtures; identify various end points that would be affected; study the mechanism of action, progression and repair; and identify (both generic and specific) that would allow the determination of the health of the organism and develop qualitative and quantitative health assessment methods so as to assess multiple health effects (Hansen et al. 1998)

Table 12.4 Toxic equivalency factors (TEFs) for human and mammalian Risk Assessment

Congener (Dioxins)	WHO TEF
2,3,7,8 – TetraCDD	1
1,2,3,7,8 – PentaCDD	1
1,2,3,4,7,8 – HexaCDD	0.1
1,2,3,6,7,8 – HexaCDD	0.1
1,2,3,7,8,9 – HexaCDD	0.1
1,2,3,4,6,7,8 HeptaCDD	0.01
OctaCDD	0.0003

CDD Chlorinated dibenzodioxin

Where contaminants share structural similarities such as Dioxins, PCBs and Polycyclic Aromatic Hydrocarbons the use of Toxic Equivalency Factors has been proposed. Different contaminants are given toxicity "scores" that are fractions of the toxicity of another in the chemical group for which there is adequate toxicity data. Given a mixture of the contaminants, a cumulative toxicity score can be determined. The IPCS has published TEFs for several dioxins and these are shown in Table 12.4.

World Health Organization (WHO) The WHO re-evaluation of human and mammalian toxic equivalency factors (TEFs) of dioxins and dioxin-like compounds (van den Berg et al. 2006).

Biological methods such as bioassays are being appraised for their application to the assessment of the toxicity of mixtures. Rodents or other mammals may be administered extracts so that a toxicity value such as an LD_{50} can be determined. These methods are expensive and time consuming which usually precludes their use in Risk Assessment. Similar techniques using aquatic species such as Daphnia are less expensive and time consuming, but are disadvantaged by the greater toxicokinetic and toxicodynamic differences between the species used and humans (Pollak 1996). There are in vitro tests such as the Microtox test and the Submitochondrial Particle Test, but these require validation for use in Risk Assessment.

Useful information for exposures to mixtures of hazards may be available from epidemiological studies of similar or closely similar mixtures.

12.5.10 Checklist for Toxicological Appraisals

The following checklist is adapted from US EPA (1995) and can be used when appraising toxicological information.

12.5.10.1 Hazard Identification

1. What is the key toxicological study (or studies) that provides the basis for health concerns?

- How good is the key study?
- Are the data from laboratory or field studies? Are the data for single species or multiple species?
- If the hazard is carcinogenic, comment on issues such as: observation of single or multiple tumour sites; occurrence of benign or malignant tumours; certain tumour types not linked to carcinogenicity; use of the maximum tolerated dose
- If the hazard is other than carcinogenic, what endpoints were observed, and what is the basis for the critical effect?
- What other studies support this finding
- What valid studies conflict with this finding.
- What are the significant data gaps?

As many relevant studies as possible should be collated and rigorously assessed as to their strengths and weaknesses to determine the key studies. This is particularly important where quantitative risk estimates will be undertaken or where there are apparently contradictory studies; in the latter case, the studies that are considered to be adequate in their design and interpretation will need to be appraised to determine the overall weight-of-evidence. See Section 2.6 for further information on weight of evidence.

2. Besides the health effect observed in the key study, are there other health endpoints of concern?
3. Consider available epidemiological or clinical data. For epidemiological studies:

   - What types of studies were used, i.e., ecologic, case-control, cohort?
   - Assess the degree to which exposures were adequately described.
   - Assess the degree to which confounding factors were adequately accounted for.
   - Assess the degree to which other causal factors were excluded.

   For further information refer to Section 3.4 "Assessing the relationship between a possible cause and an outcome".
4. How much is known about the biological mechanism by which the agent produces adverse effects?

   - Consider relevant studies on mechanisms of action which may include metabolism studies.
   - Does this information aid in the interpretation of the toxicity data?
   - What are the implications for potential health effects?

5. Consider any negative or equivocal findings in animals or humans, and whether these data were considered in the Hazard Identification.
6. Consider the Hazard Identification and consider the significance of each of the following.

   - confidence in conclusions;
   - alternative conclusions that are also supported by the data;

- significant data gaps; and
- major assumptions.

12.5.10.2 Characterisation of Dose-Response

1. What data were used to develop the dose-response curve? Would the result have been significantly different if based on a different data set?

   If animal data were used:

   - What species were used? The most sensitive, average of all species, or other?
   - Were any studies excluded? Why?

   If epidemiological data were used:

   - Which studies were used? Only positive studies, all studies, or some other combination?
   - Were any studies excluded? Why?
   - Was a meta-analysis performed to combine the epidemiological, studies? What approach was used? Were studies excluded? Why?

2. What model was used to develop the dose-response curve? What rationale supports this choice? Is chemical-specific information available to support this approach?

   For non-carcinogenic hazards:

   - How was the Tolerable Intake (or the acceptable range) estimated?
   - What assumptions or uncertainty factors were used?
   - What is the confidence in the estimates?

   For carcinogenic hazards:

   - What dose-response model was used? What is the basis for the selection of the particular dose-response model used? Are there other models that could have been used with equal plausibility and scientific validity?
   - What is the basis for selection of the model used in this instance?

3. Discuss the route and level of exposure observed in the toxicology or epidemiology studies, as compared to the expected human exposures in the situation under appraisal.

   - Are the available data from the same route of exposure as the expected human exposures? If not, are pharmacokinetic data available to extrapolate across route of exposure?
   - What is the degree of extrapolation from the observed data in the toxicological or epidemiological studies to the expected human exposures in the situation under appraisal (one to two orders of magnitude? multiple orders of magnitude)? What is the impact of such an extrapolation?

12.5.11 Uncertainty and Variability in Hazard Assessment

As with other components of Risk Assessment, the sources and magnitude of uncertainty and variability in Hazard Assessment should be identified, estimated and taken into account in decision making.

12.5.12 Sources of Toxicological and Tolerable Intake Data

There are many sources of toxicological appraisal and toxicologically-based guidance values. The nature of the guidelines may be affected by the need to address local issues or the incorporation of local science policy.

In Australia a hierarchy of sources was developed to assist risk assessors in the selection of sources of information so that their Risk Assessments would be acceptable to regulatory agencies. The hierarchy has sources grouped into "levels" which are given in the order of preference. In general, published Australian ADIs should be used, but other data may be used with appropriate justification. All documents, particularly those in the lower categories require rigorous appraisal for relevance, validity and accuracy. Other jurisdictions are likely to have lists of acceptable sources or similar hierarchies.

Acknowledgement The material in this chapter is derived from Environmental Health Risk Assessment Guidelines for Assessing Human Health Risks from Environmental Hazards (NHMRC 2004) and is used with permission. The material has been summarised by Andrew Langley for this book chapter. The principal author and editor of the original publication was Andrew Langley. Jack Dempsey and Les Davies co-wrote the Hazard Identification – Toxicology section and Roscoe Taylor and Andrew Langley co-wrote the Hazard Identification – Epidemiology section.

Qualifier As the book is about contamination, the word "contaminant" is used in this chapter and generally refers to chemical agents that are not naturally occurring but the term may also refer to physical or microbiological agents in some circumstances.

Glossary

Adapted from NHMRC (1999)

ADI Acceptable Daily Intake. The daily intake of a chemical which, during a lifetime, appears to be without appreciable risk, on the basis of all the facts known at the time. It is expressed in milligrams per kilogram of body weight per day (mg/kg_{bw}/day) (WHO 1989). For this purpose, "without appreciable risk" is taken to mean that adverse effects will not result even after a lifetime of exposure. Furthermore, for a pesticide residue, the acceptable daily intake is intended to give a guide to the maximum amount that can be taken daily in the food without appreciable risk to the consumer. Accordingly, the figure is derived as far as possible from feeding studies in animals. See also "Guidance values", "RfD" and "TDI".

Adverse Effect The change in morphology, physiology, growth, development or life span of an organism which results in impairment of functional capacity or impairment of capacity to compensate for additional stress or increase in susceptibility to the harmful effects of other environmental influences. Some adaptive changes are not generally considered to be adverse e.g., some changes in enzyme levels.

Agent Any chemical, physical, biological or social substance or factor being assessed, unless otherwise noted.

Applied Dose Amount of an agent presented to an absorption barrier and available for absorption. The amount may be the same or more than the absorbed dose.

Bias A process resulting in a tendency to produce results that differ in a systematic value from the true values. Also known as systematic error (Beaglehole et al. 1993).

BMD Benchmark Dose. The dose associated with a given incidence (e.g., 1%, 5% or 10% incidence) of effect, the Benchmark Risk, based on the best-fitting dose-response curve.

Bioavailability The ratio of the systemic dose to the applied dose.

Carcinogen Chemical, biological or physical cancer-causing agent.

Carcinogenesis The origin, causation and development of tumours. The term applies to all forms of tumours (e.g., Benign and malignant).

Carcinogenicity The ability to produce tumours, which may be benign or malignant (IEH 1999).

Chronic toxicity The ability to produce an adverse effect which persists over a long period of time, whether or not it occurs immediately upon exposure to a contaminant or is delayed, or an effect which is only induced by prolonged exposure to a contaminant (IEH 1999).

Confidence Weight assigned by the evaluator to the quality of the information available (high, medium or low confidence) to indicate that a contaminant possesses certain toxicological properties.

Confidence Limits A range of values determined by the degree of presumed random variability in a set of data, within which the value of a parameter, e.g., The mean, lies, with a specified level of confidence or probability (e.g. 95%). The confidence limit refers to the upper or lower value of the range (DOH 1991).

Confounding Factor A factor that distorts the apparent effect or magnitude of the effect of a study factor or risk. Such factors must be controlled for in order to obtain an undistorted estimate of a given effect (DOH 1991).

Critical Effect(s) The adverse effect judged to be the most important for setting an acceptable human intake or exposure. It is usually the most sensitive adverse effect, i.e., that with the lowest effect level, or sometimes a more severe effect, not necessarily having the lowest effect level (IEH 1999).

Dermal Of the skin, through or by the skin.

Dose A stated quantity or concentration of a contaminant to which an organism is exposed over a continuous or intermittent duration of exposure. It is most commonly expressed as the amount of test contaminant per unit weight of test animal (e.g., mg/kg$_{\mathrm{bw}}$).

The applied dose is the amount of contaminant in contact with the primary absorption boundaries (e.g., skin, lungs, and gastrointestinal tract) and available for absorption The absorbed dose is the amount crossing a specific absorption barrier (e.g., the exchange boundaries of skin, lung, and digestive tract) through uptake processes. The amount of the contaminant available for interaction by any particular organ or cell is termed the delivered dose of that organ or cell (EPA 1992, p. 22933). The systemic dose is the dose to which the whole, or extensive parts, of the body is exposed. The absorbed dose may not be the systemic dose as contaminants absorbed in the digestive tract may be removed by the liver and not enter the systemic circulation.

Dosage A general term comprising the dose, its frequency and the duration of dosing. Dosage is properly applied to any rate or ratio involving a dose. Dosages often involve the dimension of time (e.g., mg/kg$_{\mathrm{bw}}$/day), but the meaning is not restricted to this relationship (Hayes 1991).

Dose-response Assessment Determination of the relationship between the magnitude of the dose or level of exposure to a contaminant and the incidence or severity of the associated adverse effect (IEH 1999).

Dose-response relationship The correlative association existing between the dose administered and the response (effect) or spectrum of responses that is obtained. The concept expressed by this term is indispensable to the identification, evaluation, and interpretation of most pharmacological and toxicological responses to contaminants. The basic assumptions which underlie and support the concept are: (a) the observed response is a function of the concentration at a site, (b) the concentration at a site is a function of the dose, and (c) response and dose are causally related (Eaton and Klaassen 1996). The existence of a dose-response relationship for a particular biological or toxicological response (effect) provides a defensible conclusion that the response is a result of exposure to a known contaminant.

Endpoint An observable or measurable biological event used as an indicator of the effect of a contaminant on a biological system (cell, organism, organ etc.).

Environmental health Those aspects of human health determined by physical, chemical, biological and social factors in the environment. Environmental health practice covers the assessment, correction, control and prevention of environmental factors that can adversely affect health, as well as the enhancement of those aspects of the environment that can improve human health.

Environmental monitoring The monitoring of the concentration of contaminants in the physical environment of air, water, soil and food.

Epidemiology The study of the distribution and determinants of health related states or events in specified populations, and the application of the study to the control of health problems (Last 1988)

Expert An expert has (1) training and experience in the subject area resulting in superior knowledge in the field (2) access to relevant information, (3) an ability to process and effectively use the information, and (4) is recognised by his or her peers or those conducting the study as qualified to provide judgements about assumptions, models, and model parameters at the level of detail required (NCRP 1996).

Exposure Contact of a contaminant, physical or biological agent with the outer boundary of an organism, e.g., Inhalation, ingestion or dermal contact.

Exposure Assessment The estimation (qualitative or quantitative) of the magnitude, frequency, duration, route and extent (for example, number of organisms) of exposure to one or more contaminated media for the general population, for different subgroups of the population, or for individuals.

Exposure Route The way a contaminant enters an organism after contact e.g., By ingestion, inhalation, or dermal absorption (EPA 1992, p. 22933).

Extrapolation For dose-response curves, an estimate of the response at a point outside the range of the experimental data. Also refers to the estimation of a response in different species or by different routes than that used in the experimental study of interest.

Factor A single factor or product of several single factors used to derive an acceptable intake. These factors account for adequacy of the study, interspecies extrapolation, inter-individual variability in humans, adequacy of the overall data base, nature and extent of toxicity, public health regulatory concern and scientific uncertainty.

Gene The DNA molecule of inheritance of characteristics including susceptibility to disease.

Genotoxic Agents for which a direct activity is the alteration of the information encoded in genetic material (Butterworth 1990)

Genotoxic carcinogen A contaminant which induces tumours via a mechanism involving direct damage to DNA (IEH 1999).

Genotoxicity A broad term describing the ability to produce damage to the genetic material (DNA) of cells or organisms.

Guidance Values "Values such as concentrations in air or water, which are derived after appropriate allocation of Tolerable Intake (TI) among the possible different media of exposure. Combined exposure from all media at the guidance values over a lifetime would be expected to be without appreciable health risk. The aim of a guidance value is to provide quantitative information from Risk Assessment for risk managers to enable them to make decisions concerning the protection of human health." (WHO 1994, p. 16)

Hazard The capacity of an contaminant to produce a particular type of adverse health or environmental effect, e.g., One hazard associated with benzene is that it can cause acute myeloid leukemia; or

The disposition of a thing, a condition or a situation to produce an adverse health or environmental effect; or an event, sequence of events or combination of circumstances that could potentially have adverse consequences (adapted from ACDP 1996).

Hazard identification The identification, from animal and human studies, *in vitro* studies and structure-activity relationships, of adverse health effects associated with exposure to a contaminant (IEH 1999).

Health Health is a state of complete physical, mental and social well being and not merely the absence of disease or infirmity (WHO 1946).

Health Risk Assessment The process of estimating the potential impact of a chemical, biological, physical or social agent on a specified human population system under a specific set of conditions and for a certain timeframe.

Health Risk Management The process of evaluating alternative actions, selecting options and implementing them in response to Human Health Risk Assessment. The decision making will incorporate scientific, technological, social, economic and political information. The process requires value judgements, e.g., on the tolerability and reasonableness of costs.

Immunotoxicity The ability to produce an adverse effect on the functioning of organs and cells involved in immune competence (IEH 1999).

LD_{50} The quantity of a contaminant that, when applied directly to test organisms, via inhalation, oral or dermal exposure is estimated to be fatal to 50% of those organisms under the stated conditions of the test.

Number of microorganisms of a particular species that are fatal in 50% of the host organisms.

LED_{10} Lowest Effective Dose. The lower 95% confidence limit on a dose associated with an estimated 10% increased tumour or relevant non-tumour response (US EPA 1996).

LOEL Lowest Observed Effect Level. The lowest concentration or amount of a contaminant, found by experiment or observation, that causes alterations of morphology, functional capacity, growth, development or life span of target organisms.

WHO (1990) define it as the lowest dose of a contaminant which causes changes distinguishable from those observed in normal (control) animals.

LOAEL Lowest Observed Adverse Effect Level. The lowest concentration or amount of a contaminant, found by experiment or observation, that causes adverse alterations of morphology, functional capacity, growth, development or life span of target organisms.

Life-time Covering the average life span of an organism (e.g., 70 years for humans).

Metabolite A contaminant that is the product of biochemical alteration of the parent contaminant in an organism.

Model A mathematical representation of a biological system intended to mimic the behaviour of the real system, allowing descriptions from empirical data and predictions about untested states of the system.

NOAEL The No-Observed-Adverse-Effect Level is the highest dose of a contaminant at which no toxic (i.e., adverse) effects are observed (WHO 1990). It may also be worded in more detail thus: The NOAEL is defined as the highest exposure at which there is no statistically- or biologically-significant increase in the frequency of an adverse effect when compared to a control group (National Academy of Sciences/National Research Council 1994). The definition of NOEL is equivalent, but with the removal of the term, "adverse". Often, the difficult issue in the use of the terms NOEL or NOAEL is in deciding whether a contaminant-related effect noted in a particular study is necessarily an "adverse" effect. Alterations of morphology, functional capacity, growth, development or life span of the target organism may be detected which are judged not to be adverse.

Nongenotoxic carcinogen A contaminant which induces tumours via a mechanism which does not involve direct damage to DNA (IEH 1999).

Physiologically-based pharmacokinetics (PBPK) Modelling the dose or degree of exposure to a contaminant at a target tissue, cell or receptor, by integration of pharmacokinetic data with anatomical, physiological and biochemical data (IEH 1999).

NOEL The "No-Observed-Effect Level" or "No-Observable-Effect Level" (NOEL) is the highest dose of a contaminant administered to a group of experimental animals at which there is an absence of observable effects on morphology, functional capacity, growth, development or life span, which are observed or measured at higher dose levels used in the study. Thus, dosing animals at the NOEL should not produce any biologically significant differences between the group of chemically exposed animals and an unexposed control group of animals maintained under identical conditions. The NOEL is expressed in milligrams of contaminant per kilogram of body weight per day (mg/kg$_{\mathrm{bw}}$/day) or, in a feeding study, in ppm in food (converted to mg/kg$_{\mathrm{bw}}$ of contaminant intake by measured or estimated food intake over the period of the study) The NOEL has been simply defined as the highest dose of a contaminant which causes no changes distinguishable from those observed in normal (control) animals (WHO 1990).

PTWI Provisional Tolerable Weekly Intake. The tolerable intake of a contaminant expressed as a weekly amount. The term was established by WHO (1972) for several heavy metals which "are able to accumulate within the body at a rate and to an extent determined by the level of intake and by the contaminant form of the heavy metal present in food." (WHO 1989)

Public Health The science and art of preventing disease, prolonging life and promoting health through organised efforts of society.

Reproductive toxicity The ability to produce an adverse effect on any aspect of reproductive capacity, function or outcome. It includes effects on the embryo, fetus, neonate and prepubertal organism and on adult reproductive and neuroendocrine systems (IEH 1999).

RfD Reference Dose (RfD). An estimate (with uncertainty factors spanning perhaps an order of magnitude) of the daily exposure (mg/kg$_{\mathrm{bw}}$/day) to the general human population (including sensitive sub-groups) that is likely to be without an appreciable risk of deleterious effects during a life time of exposure. It is derived from the NOAEL or the LOAEL by application of uncertainty factors that reflect various types of data used to estimate *RfD* and an additional modifying factor, which is based on professional judgement of the entire data base of the contaminant (IRIS 1996). Usually doses less than the RfD are not likely to be assoicated with adverse health risks, and are therefore less likely to be of regulatory concern. As the frequency and/or magnitude of the exposures exceeding the RfD increase, the probability of adverse effects in a human population increases. However, all doses below the RfD are not assumed to be "acceptable" (or risk-free) and nor are all doses that exceed the RfD necessarily "unacceptable" (i.e., Result in adverse effects)

Risk The probability that, in a certain timeframe, an adverse outcome will occur in a person, group of people, plants, animals and/or the ecology of a specified area that is exposed to a particular dose or concentration of a contaminant, i.e., it depends on both the level of toxicity of the contaminant and the level of exposure.

Risk Assessment The process of estimating the potential impact of a chemical, physical, microbiological or psychosocial hazard on a specified human population or ecological system under a specific set of conditions and for a certain timeframe

Risk management The process of evaluating alternative actions, selecting options and implementing them in response to Risk Assessments. The decision making will incorporate scientific, technological, social, economic and political information. The process requires value judgements, e.g. on the tolerability and reasonableness of costs.

Safety factor See factor. Safety factor usually refers to health-related concerns.

Stochastic A random probabilistic phenomenon

Teratogenicity The ability to produce a structural malformation or defect in an embryo or fetus (IEH 1999)

Threshold The lowest dose or exposure level which will produce a toxic effect and below which no toxicity is observed (IEH 1999).

Threshold Dose The lowest dose which produces an effect and below which no biological effect occurs. The acceptability and usefulness of the concept of the experimental NOEL/NOAEL depends on the scientific rationale supporting the existence and demonstrability of a threshold for responses produced by biologically active contaminant. As used here, the term "threshold" designates that level of a stimulus which comes just within the limits of perception, and below which level a recognisable response is not elicited.

TDI Tolerable Daily Intake. An estimate of the intake of a contaminant which can occur over a lifetime without appreciable health risk. It may have different units depending on the route of administration (WHO 1994). (Imray and Langley 1996, p. 18). The term, "acceptable" daily intake is used for contaminants such as pesticides (herbicides, insecticides, antifungals etc.) which are deliberately used on food crops or food-producing animals and for which some level of residues may be expected to occur in food. The term "tolerable" daily intake is used when the contaminant is a potential food or environmental contaminant. Whilst exposure should not occur, a TDI is an established health limit below which lifetime exposure should not have any adverse health effects."Acceptable Daily Intake" and "Reference Dose".

TWI Tolerable Weekly Intake. The TI expressed as a weekly amount

Tolerable Intake "an estimate of the intake of a contaminant that over a lifetime is without appreciable health risk." (WHO 1994). Examples are the ADI, TDI and Reference Dose.

Toxicity The quality or degree of being poisonous or harmful to plant, animal or human life.

Tumour A mass of abnormal, disorganised cells, arising from pre-existing tissue, which is characterised by excessive and uncoordinated cell proliferation or growth and by abnormal differentiation (specialisation). There are two types of tumours, benign and malignant. Benign tumours morphologically resemble their tissue of origin, grow slowly (may also stop growing) and form encapsulated masses; they do not infiltrate other tissues, they do not metastasise and are rarely fatal. Malignant tumours resemble their parent tissue less closely and are composed of increasingly abnormal cells genetically, morphologically and functionally. Most grow rapidly, spread progressively through adjacent tissues and metastasise to distant tissues.

Uncertainty The lack of knowledge about the correct value, e.g. a specific exposure measure or estimate

Uncertainty factor A numerical factor applied to the no-effect level to derive an exposure level considered to be without appreciable risk to health (the NEL is divided by the uncertainty factor). The magnitude of the uncertainty factor depends on the nature of the toxicity observed, the quality of the toxicological data available, and whether the effects were observed in humans or animals (IEH 1999).

Variability Measurable factors that differ e.g. height is variable across populations. The major types of variability are temporal, spatial and interindividual. They may be discrete (e.g. albinism) or continuous (e.g., body weight). It may be readily identifiable (e.g., presence of albinism) or difficult to identify (e.g., ability to detoxify a particular contaminant metabolite)

References

Advisory Committee on Dangerous Pathogens (ACDP) (1996) Microbiological Risk Assessment: an interim report. HMSO, London

Agency for Toxic Substances and Disease Registry (ATSDR) (1996) *Guidance for ATSDR Health Studies*. Division of Health Studies, ATSDR, Washington

Ames BN, Durston WE, Yamasaki E, Lee FD (1973) Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. Proc Natl Acad Sci USA 70:2281–2285

Ashby J, Tennant RW (1991) Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the US NTP. Mutat Res 257:229–306

Barnes DG, Dourson ML (1988) Reference Dose (R*f*D): description and use in health risk assessments. Reg Tox Pharmacol 8:471–486

Beaglehole R, Bonita R, Kjellstrom T (1993) Basic epidemiology. World Health Organization, Geneva

Bishop JM (1991) Molecular themes in oncogenesis. Cell 64:235–248

Boobis AR, Samuel M, Cohen SM, Dellarco V et al (2006) IPCS framework for analyzing the relevance of a cancer mode of action for humans. Crit Rev Toxicol 36(10):781–792

Bonita R Beaglehole, Kjellstrom T (2006) Basic epidemiology, 2nd edn. World Health Organization, Geneva

Breslow NE, Day NE (1987) Statistical methods in cancer research. Volume II – the design and analysis of cohort Studies. International Agency for Research on Cancer, Scientific Publication 82, Lyon

Butterworth BE (1990) Consideration of both genotoxic and non-genotoxic mechanisms in predicting carcinogenic potential. Mutat Res 239:117–132

Calabrese EJ, Baldwin LA (1994) Improved method of selection of the NOAEL. Reg Tox Pharmacol 19:48–50

Calabrese EJ, Gilbert RO (1993) Lack of total independence of uncertainty factors(UFs): implication for size of the total uncertainty factor. Reg Tox Pharmacol 17:44–51

Crump K (1984) A new method for determining allowable daily intakes. Fundam Appl Toxicol 4:854–891

Crump K (1985) Mechanisms leading to dose-response models. In: Ricci P (ed) Principles of health risk assessment, Prentice Hall, Englewood Cliffs, NJ

Department of Health (DOH) (1991) Guidelines for the evaluation of chemicals for carcinogenicity. Committee on carcinogenicity of chemicals in food, consumer products and the environment. Department of Health, Report RHSS 42, HMSO, London

Dickens A, Robinson J (1996) Statistical approaches. In: Evans GO (ed) Animal clinical chemistry: a primer for toxicologists. Taylor and Francis, London, pp 45–48

Doull J (1980) Factors influencing toxicology. In: Casarett and Doull's toxicology, 2nd edn. Macmillan, New York, pp 70–83

Dourson et al (1996) Evolution of science-based uncertainty factors in noncancer risk assessment. Reg Toxicol Pharmacol 24(2):121–125

Dybing E (2002) Development and implementation of the IPCS conceptual framework for evaluating mode of action of chemical carcinogens. Toxicology 181–182: 121–125

Eaton DL, Klaassen D (1996) Principles of toxicology. In: Casarett and Doull's toxicology; The basic science of poisons, 5th edn. McGraw-Hill, New York, pp 13–33

Elliott P, Cuzick, English D, Stern R (1992) Geographical and environmental epidemiology: methods for small-area studies. World Health Organization, Oxford

Fitzgerald DJ (1993) Carcinogenic soil contaminants: an Australian approach. In: Langley AJ, van Alphen M (eds) The health risk assessment and management of contaminated sites – proceedings of the second national workshop on the health risk assessment and management of contaminated sites. South Australian Health Commission, Adelaide, pp 51–64

Fowle JR, Alexeeff GV, Dodge D (1999) The use of benchmark dose methodology with acute inhalation lethality data. Reg Toxicol Pharmacol 29:262–278

Gad SC, Weil CS (1986) Statistics and experimental design for toxicologists. Telford, NJ

Gad SC, Weil CS (1989) Statistics for toxicologists. In: Hayes AW (ed) Principles and methods of toxicology, 3rd edn. Raven, New York, Chapter 7, pp 221–274

Gaylor DW, Gold LS (1998) Regulatory cancer risk assessment based on a quick estimate of a benchmark dose derived from the maximum tolerated dose. Erg Toxic Pharmacy 28:222–225

Gaylor DW, Ryan L, Crews D, Zhu Y (1998) Procedures for calculating benchmark doses for health risk assessment. Erg Toxic Pharmacy 28:150–164

Greenhalgh T (1997) Papers that summarise other papers – systematic reviews and meta-analysis. Br Med J 315:672–675

Gregg MB (ed) (1996) Field epidemiology. Oxford University Press, New York

Hansen H, De Rosa CT, Pohl H, Fay M, Mumtaz MM (1998) Public health challenges posed by chemical mixtures. Environ Health Perspect 106(Suppl 6):1271–1280

Haseman JK, Huff J, Boorman GA (1984) Use of historical control data in carcinogenicity studies in rodents. Toxicol Pathol 12:126-135

Hayes WJ (1991) Dosage and other factors influencing toxicology. In: Hayes WJ, Laws ER (eds) Handbook of toxicology, vol 1, general principles. Academic, San Diego, pp 40–41

Health Council of Netherlands (1994) Risk assessment of carcinogenic chemicals in the netherlands. Regul Toxicol Pharmacol 19:14–30

Hellard M (1999) The water quality study approaches a conclusion. Water Quality News, Issue 10:2

Heywood R (1981) Target organ toxicity. Toxicol Lett 8: 349-358

Hill AB (1965) The environment and disease: association or causation? Proc Royal Soc Med 58:295–300

International Agency for Research on Cancer (IARC) (1978) Some N-nitroso compounds vol 17. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. IARC, Lyon

IARC (1983) Miscellaneous pesticides vol 30. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. IARC, Lyon

IARC (1987) Overall evaluations of carcinogenicity: an updating of IARC monographs volumes 1 to 42 – suppl. 7. IARC monographs on the evaluation of carcinogenic risks to humans. IARC, Lyon

IARC (1999) The use of short- and medium-term tests for carcinogens and data on genetic effects in carcinogenic hazard evaluation no. 146. International Agency for Research on Cancer Scientific, IARC, Lyon

Imray P, Langley A (1996/1998). Health-based soil investigation levels. National environmental health forum monographs. Soil series no. 1. South Australian Health Commission, Adelaide, 1996

IEH (1999) Risk assessment approaches used by UK government for evaluating human health effects of chemicals. Institute for Environment and Health, Leicester

IRIS (1996) Integrated risk information system. US Environmental Protection Agency, Washington, DC

Johannsen FR (1990) Risk Assessment of carcinogenic and noncarcinogenic chemicals. Crit Rev Toxicol 20(5):341–367

Klaassen CD (ed) (1996) Casarett and Doull's toxicology; The basic science of poisons, 5th edn. McGraw-Hill, New York, p 4

Korach KS (ed) (1998) Reproductive and developmental toxicology. Marcel Dekker, New York

Last JM (1988) A dictionary of epidemiology, 2nd edn. Oxford University, Oxford

Lee PN (1993) Statistics. In: Anderson D, Conning DM (eds) Experimental toxicology: the basic issues. Royal Society of Chemistry, Cambridge, UK

Lewis SC, Lynch JR, Nikiforov AI (1990) A new approach to deriving community exposure guidelines for 'no observed adverse effect levels'. Reg Tox Pharm 3:17–34

Lilienfield AM, Lilienfield DE (1980) Foundations of epidemiology, 2nd edn. Oxford University Press, New York

Lovell DP, Thomas G (1996) Quantitative risk assessment and the limitations of the linearised multistage model. Human Exp Toxicol 15:87–104

Lu FC (1995) A review of the acceptable daily intakes of pesticides assessed by WHO. Regul Toxicol Pharmacol 21:352–364

Lu FC, Sielken RL Jr (1991) Assessment of safety/risk of chemicals: inception and evolution of the ADI and dose-response modelling procedures. Toxicol Lett 59:5–40

McClellan RO, Henderson RF (eds) (1989) Concepts in inhalation toxicology. Hemisphere Publishing, New York, NY

McMichael AJ (1991) Setting environmental exposure standards: current concepts and controversies. Int J Environ Health Res 1:2–13

Moolgavkar SH, Moller H, Woodward A (1999) Principles of the epidemiological approach to QEP. In: Moolgavkar SH, Krewski L, Zeise L et al (eds) Quantitative estimation and prediction of human cancer risks. International Agency for Research on Cancer, Scientific Publication 131, Lyon

Mundt KA, Tritschler JP, Dell LD (1998) Validity of epidemiological data in risk assessment applications. Human Ecol Risk Assess 4:675–683

NHMRC (1999) Toxicity assessment for carcinogenic soil contaminants Canberra: national health & medical research council. Previously titled Draft cancer risk assessment for environmental contaminants. Commonwealth of Australia, Canberra

NHMRC (2000) How to review the evidence: systematic identification and review of the scientific literature. Commonwealth of Australia, Canberra

NHMRC (2004) Environmental health risk assessment guidelines for assessing human health risks from environmental hazards. Commonwealth of Australia, Canberra. Available at: www.health.gov.au/internet/main/Publishing.nsf/Content/ohp-ehra-2004.htm

NHMRC (2006) Ambient air quality standards setting: an approach to health-based hazard assessment. Commonwealth of Australia, Canberra

National Committee on Radiation Programs (NCRP) (1996) NRCP commentary no. 14. A guide for uncertainty analysis in dose and risk assessments related to scientific contamination. National Committee on Radiation Programs, Washington

Nurminen M (1995) Linkage failures in ecological studies. World Health Stat Quart 48: 78–84

Organisation for Economic Cooperation and Development (OECD) (1982) Good laboratory practice in the testing of chemicals. Organisation for Economic Cooperation and Development, Paris, 62 pp

OECD (1998) OECD guidelines for the testing of chemicals. Section 4: health effects. vol 2, 10th addendum, October 1998. Organisation for economic cooperation and development, Paris

OECD (Various Dates) OECD guidelines for the testing of chemicals. Section 4: health effects. Organisation for economic cooperation and development, Paris

Paustenbach DJ (1995) The practice of health risk assessment in the United States (1975-1995): how the US and other countries can benefit from that experience. Human Ecol Risk Assess 1:29-79

Paynter OE (1984) Oncogenic potential: guidance for analysis and evaluation of long term rodent studies. Evaluation procedure #1000.1 office of pesticide and toxic substances, EPA, Washington, DC

Pershagen G (1999) Research priorities in environmental health. BMJ 318:1636–1637

Pollak JK (1996) Free radical biochemistry of contaminants. In: Langley A, Imray P, Lock W, Hill H (eds) The health risk assessment and management of contaminated sites, Proceedings of the fourth national workshop on the health risk assessment and management of contaminated sites. Contaminated sites monograph series number 7. South Australian Health Commission, Adelaide, pp 323–338

Porta M (2008) A dictionary of epidemiology, 5th edn. Oxford University, New York

Purchase IFH, Auton TR (1995) Thresholds in chemical carcinogenesis. Regulat Toxicol Pharmacol 22:199–205

Renwick AG, Walker R (1993) An analysis of the risk of exceeding acceptable or tolerable daily intake. Regulat Toxicol Pharmacol 18:463–480

Roseman JM (1998) What can be done to ensure the usefulness of epidemiological data for risk assessment purposes? Human Ecol Risk Assess 4:737–746

Rothman KJ, Greenland S (1997) Causation and causal inference. In: Detels et al (eds) Oxford textbook of pubic health, 3rd edn. Oxford Medical, New York

Roubicek CB, Pahnish OF, Taylor RL (1964) Growth of rats at two temperatures. Growth 28: 157–164

Samet JM, Schnatter R, Gibb H (1998) Invited commentary: epidemiology and risk assessment. Am J Epidemiol 148:929–936

Saunders LD, Chen W, Hrudey SE (1997) Studies of hazardous waste sites and human health: evaluation of internal validity. Environ Rev 5:167–180

Seefeld MD, Peterson RE (1984) Digestible energy and efficiency of feed utilisation in rats treated with 2,3,7,8,-tetrachlorodibenzo-p-dioxin. Tox Appl Pharmacol 74:214-222

Swaim LD, Taylor HW, Jersey GC (1985) The effect of handling techniques on serum ALT activity in mice. J Appl Toxicol 5:160–162

Task Force of Past Presidents (1982) Animal data in hazard evaluation: paths and pitfalls. Toxicol Sci 2:101–107

Tallarida RJ, Murray RB (1987) Manual of pharmacologic calculations with computer programs, 2nd edn. Springer, New York

Thomas SP, Hrudey SE (1997) Risk of death in Canada: what we know and how we know it. University of Alberta, Edmonton

Thompson GE, Scheel LD, Lowe D (1982) Seasonal alteration in response to stress or physiological change. Drug Chem Toxicol 5:189–199

Tucker MJ (1979) The effect of long-term food restriction on tumours in rodents. Int J Cancer 23:803–807

Tyl RW, Marr MC (1997) In: Hood RD (ed) Handbook of developmental toxicology. CRC, Boca Raton

US Environmental Protection Agency (US EPA) (1983) Good laboratory practice standards: toxicology testing. Fed Reg 48: page 53922, and: pesticide programs: good laboratory practice standards. final rule. Fed Reg 46 (No. 230): 53946–53969, Tuesday 11/29/83

US EPA (1986) The risk assessment guidelines of 1986. Office of Health and Environmental Assessment, EPA/600/8 87/045. United States Environmental Protection Agency, Washington DC

US EPA (1989) Risk assessment guidance for Superfund, volume 1. Human health evaluation manual, (Part A). Interim final. EPA/540/1–89/002. US Environmental Protection Agency, Office of Emergency and Remedial Response, Washington DC

US EPA (1992) Guidelines for exposure assessment; notice. Part VI. federal register 29 May 1992 57(104):22888–22938

US EPA (1995) Guidance for risk characterization. US Environmental Protection Agency Science Policy Council. Washington, DC

US EPA (1996) Draft Revision to the guidelines for carcinogen risk assessment. Office of Health and Environmental Assessment, Office of Research and Development, United States Environmental Protection Agency, Washington, DC

US EPA (2005) Guidelines for carcinogen risk assessment. Office of Health and Environmental Assessment, Office of Research and Development, United States Environmental Protection Agency, Washington, DC

Vainio H, Wilbourn J (1992) Identification of carcinogens within the IARC Monograph program. Scand. J Work Environ Health 18(Suppl 1):64–73

Vainio H, Heseltine E, McGregor D, Tomatis L, Wilbourn J (1992) Working group on mechanisms of carcinogenesis and the evaluation of carcinogenic risks. Cancer Res 52:2357–2361

Van den Berg et al (2006) The WHO re-evaluation of human and mammalian toxic equivalency factors (TEFs) of dioxins and dioxin-like compounds. 93(2):223–41. Epub 2006 Jul 7. Available at http://www.who.int/ipcs/assessment/tef_update/en. Accessed on 25 August 2010

Waner T (1992) Current statistical approaches to clinical pathology data from toxicological studies. Toxicol Pathol 20(3):Part 2

Weil SC, McCollister DD (1963) Relationship between short-and-long term feeding studies in designing an effective toxicity test. Agr Food Chem 11:486–491

Whysner J, Williams GM (1992) International cancer risk assessment: the impact of biologic mechanisms. Regul Toxicol Pharmacol 15:41–50

World Health Organization (1946) Preamble to the constitution of the world health organization as adopted by the international health conference, New York, pp 19–22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948

World Health Organization (WHO) (1972) Evaluation of certain food additives and the contaminants mercury, lead, and cadmium. Sixteenth report of the Joint FAO/WHO Expert Committee on Food Additives (JECFA). World Health Organization Technical Report Series 505. World Health Organization, Geneva

WHO (1978) Principles and methods for evaluating the toxicity of chemicals. Environmental health criteria 6. International Programme on Chemical Safety/World Health Organization. Geneva

WHO (1987) Principles for the safety assessment of food additives and contaminants in food. Environmental Health Criteria 70, IPCS/WHO, Geneva

WHO (1989) Evaluation of certain food additives and contaminants. Thirty third report of the Joint FAO/WHO Expert Committee on Food Additives (JECFA). World Health Organization Technical Report Series 776. World Health Organisation, Geneva

WHO (1990) Principles for the toxicological assessment of pesticide residues in food, Environmental Health Criteria 104, IPCS/WHO, Geneva

WHO (1992) Quality management for chemical safety testing. Environmental health criteria 141. International Programme on Chemical Safety/World Health Organization. Geneva

WHO (1993) Guidelines for drinking water quality. Volume 1 Recommendations, 2nd edn, WHO, Geneva

WHO (1994) Assessing human health risks of chemicals: derivation of guidance values for health based exposure limits. Environmental Health Criteria 170. IPCS/WHO, Geneva

WHO (1999) Chemical risk assessment. Training Module No. 3 WHO/PCS/99.2. World Health Organization, Geneva

WHO (2000) Evaluation and use of epidemiological evidence for environmental health risk assessment. World Health Organization Regional Office for Europe, Copenhagen

WHO (2009) Principles for modelling dose-response for the risk assessment of chemicals, Environmental Health Criteria 239, IPCS/WHO, Geneva

Yassi A, Kjellstrom T, de Kok T, Guidotti T (2001) Basic environmental health. Oxford University, New York

Zbinden G (1979) The role of pathology in toxicity testing. In: Progress in toxicology, vol 2. Springer, New York, pp 8-18

Part IV
Ecological Aspects

Chapter 13
Introduction to Ecological Risk Assessment

Frank A. Swartjes, Anton M. Breure, and Michel Beaulieu

Abstract The topsoil is the most biologically diverse part of the earth, harbouring more than one billion organisms per square meter. These soil organisms live in extremely complex mutual interaction and, additionally, in similarly complex interactions with their physical and chemical environment. Although not always acknowledged by the general public, the soil ecosystems perform so-called Ecosystem Services which are very important for society. Some of these Ecosystem Services, described in detail in this chapter, are soil structuring, humus formation, nutrient supply, cleaning function, disease control, and – only recently recognised – energy-related processes. The conclusion to be drawn is that intensive communication about Ecological Risk Assessment is a necessity, both to guarantee that appropriate ecological protection is on the political agenda and to justify protection of the soil ecosystem and the costs involved for the tax payer. Soil contamination has a big impact on the soil ecosystem. Ecological Risk Assessment is an extremely useful process for supporting the decisions taken concerning contaminated sites. The general target for Ecological Risk Assessment is Ecological Health (the preferred state) rather than the Ecological Integrity (the unimpaired condition), and this ideally at the level of a whole ecosystem. The important factors that relate to ecological effects in soil will be introduced in this chapter, factors such as bioavailability, food supply, sealing, resilience and recovery, adaptation, land use, secondary poisoning, the food web approach, wildlife protection, scale and contaminant pattern, and spatial planning. Finally, insight will be provided as to how Ecological Risk Assessment actually works in practice.

F.A. Swartjes (✉)
National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: frank.swartjes@rivm.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_13,

Contents

13.1 Introduction

13.1.1 Vital Soil

The biggest mistake that could be made regarding the vitality of soil is to consider it as ‘dead material’. Slightly more realistic, many people believe that soil is inhabited by only a handful of organisms that they can see with the naked eye, mainly moles, earthworms and maybe ‘some obscure beetles’. In reality, soil is the habitat of an enormous number of organisms. In fact, the topsoil is the part of the earth with the highest Biodiversity (Takeda et al. 2007), which includes more than one fourth of all living species on earth (European Commission DG ENV, 2010). It is buzzing under our feet even if most of us ignore it. According to the *Encyclopaedia Britannica* one square meter of rich soil can harbour as many as 1,000,000,000 organisms. Thus, most of the people that are not directly involved in soil biology or ecology are impressed when they learn about the huge number of species and organisms in soil. They are even more impressed when they find out what important functions the soil ecosystem performs and when they are confronted with the complexity of the way organisms are performing these functions. Just the partition between the different types of organisms and the succession of tasks performed by them surpasses the human imagination.

It is generally acknowledged that the strain on the environment, and on the soil ecosystems, is enormous. In many countries protection of the soil ecology is the second most important protection target in dealing with contaminated sites after human health protection. However, since the very beginning of soil policy in the late 1970s only a few countries have given the same weight to ecological soil protection as to human health protection. The countries that have ecological protection included in their soil policies often claim that both human health and ecology deserve unconditional protection in a civilized society. Appreciation of soil ecology rests on the two following basic principles which are closely interrelated:

- The optimisation of *biodiversity*: the most basic principle, on which the appreciation of the soil ecosystem is based, is the protection of living beings (soil organisms), because of their intrinsic value and independent of their possible use.
- The so-called *Ecosystem Services*, these are the processes performed by organisms within ecosystems, processes that are essential for the survival of mankind. This relates to the role of soil organisms in profitable functions such as the growth of plants and trees, the degradation of organic matter, the purification of soil from contaminants, the protection of groundwater quality, and the role played in the structure formation of soil that is important, for example, for agriculture and water regulation. All these processes are indispensable for agriculture and nature and, hence, essential for the quality of human life.

The subject of Ecosystem Services is far less known to the public than the subject of Biodiversity and, hence, almost exclusively appreciated (but highly appreciated!) by biological and ecological experts and other insiders. Recently these Ecosystem Services have received explicit attention in the EU Soil Thematic Strategy. Another reason to include ecological soil quality in procedures and regulations for contaminated site management is the assumption that, unlike human beings, the ecosystem cannot, or can only to a limited extent, protect itself. Humans have a voice, can evaluate the risk they are experiencing, and can contribute to solutions or even escape the problems associated with contaminated sites. Most soil organisms only have indirect possibilities, such as adaptation and avoidance, in order to escape the consequences of soil and groundwater contamination.

13.1.2 Terminology, Ranking and Classification

Soil organisms are part of soil *ecosystems*. Ecosystems may be defined as systems where chemical substances and energy are exchanged between organisms, and between organisms and their abiotic environment. In this definition the soil ecosystem includes the soil and groundwater itself. The *biocoenose* (or biocenose, biotic community, or ecological community) comprises the interactions between organisms in a specific habitat. The range of variation in the physical and chemical factors in soil is dependent on climate, soil type and land use. As a consequence, every single ecosystem is unique. Therefore, an infinite number of different ecosystems exist worldwide.

The unicity of soil ecosystems also means that different requirements apply for optimal functioning of different soil ecosystems. And since the physical and chemical factors are changing over time (for example: temperature, rain), the soil organisms must be able to perform and to survive under changing conditions. Moreover, species will adapt to changing conditions over time and, hence, the ecosystem will be changing continuously.

In the framework of this book, the protection target 'soil ecosystem' is approached from the perspective of the living organisms in soil, or in other words, the *soil biology*. According to this viewpoint, soil organisms have the highest ranking within the hierarchy of an ecosystem. Abiotic factors in soil are approached as elements that impact these organisms.

Organisms can be classified (taxonomically ranked) in different entities, for which the lowest rank is *species* (see Fig. 13.1). In total, eight higher taxonomic ranks are distinguished in biological classification, as presented in Fig. 13.1.

Several classifications exist for ecological functions (Ecosystem Services). Wallace (2007), however, claims that systems employed mix processes (means) for achieving services and the services themselves (ends) within the same classification category.

The condition of the soil ecosystem is often described as *Ecological Health* (aka: *Ecosystem Health, Soil Health*, or *soil vitality*) or *Ecological Integrity*. Ecological Integrity refers to the unimpaired condition in which ecosystems show no influence from human activities. Ecological Health relates to the preferred state of the soil ecosystem, allowing human activities to influence the state of the ecosystem.

There has been a lot of debate about the good and bad points of the phrase 'Ecological Health', especially with regard to the position of the soil ecosystem

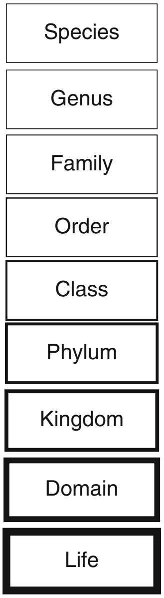


Fig. 13.1 Taxonomic ranks of organisms, including soil organisms

on the political agenda (Lackey 2001). It is often claimed that optimal Ecological Health is a broad societal aspiration rather than a concrete policy goal (e.g., Rapport 1995). Lancaster (2000) even claimed that the notion that the Ecological Health of the environment can be assessed is a ridiculous notion in a scientific context. The reasons for this are that there is no objective definition of Ecological Health and there are no methods for defining degrees of Ecological Health. The authors stated that environmental monitoring programs need to adopt a more holistic, ecosystems approach than has been used hitherto. Some claim that the health metaphor is misleading (e.g., Kapustka and Landis 1998), while others protest about the lack of clear definition (e.g., Callicott et al. 1999). One of the advantages of the phrase is that it is a straightforward, intuitive metaphor (De Leo and Levin 1997). Since it is associated with human health – an indisputably important, if not the most important issue in life – the phrase has a positive connotation and is generally well received by the general public.

From a more general perspective, the phrase *soil quality* is often used as a holistic quality mark that includes the condition of the soil ecosystem.

Processes are performed by organisms with distinct roles in ecosystems in order to maintain those ecosystems, and this determines the type of ecosystem. These functions may have different names such as Functional services, Functional (bio)diversity, and Life support functions. In this book, these functions are called *Ecosystem Services*, since the functions may be seen as being beneficial for mankind. This is analogous to the Millennium Ecosystem Assessment (Millennium Ecosystem Assessment 2005).

13.1.3 Public Perception

Although the appreciation of nature and wildlife have experienced an enormous increase in terms of interest over the last few decades, public appreciation of protection of living beings *in soil* has generally been limited. One important aspect that influences the public's appreciation is that nobody has ever seen the entity 'ecosystem', let alone the soil ecosystem. Nevertheless, the ubiquity of the word *ecology* or *ecosystem*, almost explicitly with a positive connotation, is impressive. The words are used in the advertisement of widely differing goods, the promotion of financial products and in school agendas. The term is even found in a wide variety of contexts on fashionable T-shirts.

The most extreme negative opinion of the soil ecosystem relates to the *appearance* of organisms in soil: 'One cannot see them and if you could see them you would not find them pretty.' The most extreme of the positive points of view about soil organisms comes from individuals who claim that all living matter must be able to count on unconditional protection, independent of appearance or practical use. This viewpoint is sometimes based on religious or spiritual principles (Gottlieb 2006). In the UN convention of Rio, the first and most important reason for protection of Biodiversity was its *intrinsic value* (UNCED 1992), which is an ethical reason. A Forum on Religion & Ecology has been active since 1996 with research, education and outreach regarding the relationship between ecology and religion (Religion and Ecology 2009). The forum states that, in spite of the abundant scientific knowledge about ecology, widespread global ecological degradation continues. Therefore, they believe that religious, spiritual and ethical awareness is crucial for the support of a vibrant community on earth.

However, just like the soil compartments as a whole, see Section 1.1.3, *soil organisms* are not among the most popular organisms. Almost everybody is aware of the fact that soils are occupied by earthworms and moles. But there are many more organisms that can be seen with the naked eye, for example, sow bugs, millipedes, centipedes, slugs, snails, and springtails (Sullivan 2004). However, not many kids will mention a nematode or an earthworm as their favourite pet. Another group (kingdom) of well-known organisms that lives on the soil surface and in the first few centimetres of soil is that of the arthropods. This group includes *insects* such as beetles and ants, *arachnids* such as spiders and mites, *chilopods* such as centipedes, and *diplopods* such as millipedes. Unfortunately, these organisms are generally considered rather more as a nuisance to humans than as useful, and are certainly not considered popular organisms.

Although their functions are unknown to a large section of the human population, and the majority of the organisms in soil is not visible, one typically has to learn what is behind the soil ecosystem to appreciate it. In this regard, enthusiastic biology teachers may provide children with knowledge they can benefit from all through their lives.

One element that can count on a huge appreciation by many individuals is the above-ground biology. Human life is unimaginable without plants, trees and animals. Men depend on them for food and they are important parts of the environment as it is experienced, certainly in rural areas and nature reserves, but also in urban areas and even in big cities, in parks and even in busy streets lined with trees. Ironically, the general public is not always aware of the fact that above-ground biology would not have much of a chance without the amazing activities of enormous numbers of organisms in the soil, and that lie underneath the soil surface.

In conclusion, scientists such as eco(toxico)logists, biologists and soil scientists have the serious duty of communicating the blessings of soil ecology to decision-makers, regulators and consultants. Teachers and scientific journalists can play an important role in communicating the assets of soil ecology to the general public.

13.2 Soil Biology

13.2.1 Soil Life

Millions of millions of organisms live in each handful of soil, differing widely in appearance and behaviour. These organisms include bacteria, algae, fungi, protozoa, and larger and more complex organisms, such as nematodes, micro-arthropods, potworms, earthworms, insects, small vertebrates, and (the roots of) plants. These soil organisms live in an extremely complex mutual interaction, and likewise in complex interactions with their physical and chemical environment.

The activity in soils is impressive. In an arable soil in Western Europe, the microbial community decomposes about 5,000 kg of carbon per hectare and per year, and 100 kg of nitrogen (N) per hectare and per year (Bloem and Breure 2003). Soil organisms perform an extremely wide range of processes which mainly serve the food supply of the organisms. Given the complexity of the mutual interactions of organisms in soil, and their interactions with the physical and chemical environment, it is nearly impossible to comprehend, let alone describe, the complete functioning of the whole ecosystem.

A relatively simple way to illustrate interactions between organisms in soil is by the food web, that is, a diagram that illustrates the feeding relationships between organisms. For a simplified example, see Fig. 13.2 (Ingham 2000a).

13.2.2 Classification of Organisms

13.2.2.1 Types of Classification

Since so many types of organisms occupy the soil, a classification of these organisms is essential. Because of the enormous diversity of soil organisms, several criteria for classification can be used and several classifications do exist. For example, the ecosystem can be subdivided into several functional groups, which are each responsible for a specific function, for example, protein-degrading bacteria, or humus-degrading actinomycetes. Alternatively, soil organisms can be categorized into different classes, according to size:

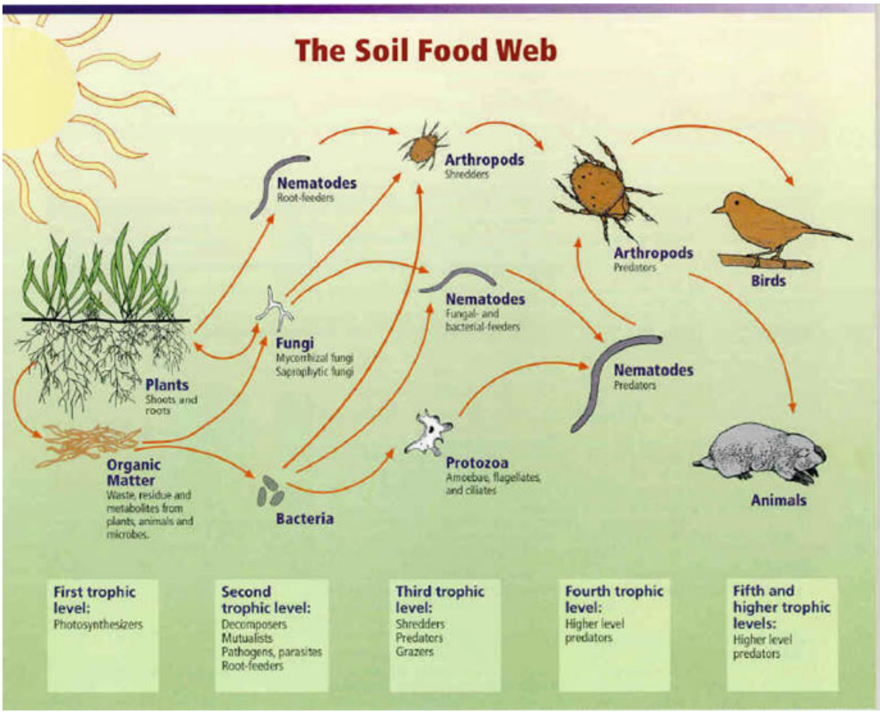


Fig. 13.2 A simplified example of a food web, illustrating the feeding relationship between organisms (source: Ingham 2000a; reproduced with permission)

- Megafauna: organisms larger than 20 mm, for example, moles, rabbits and rodents.
- Macrofauna: organisms with a size of between 2 and 20 mm, for example, woodlice, earthworms and beetles.
- Mesofauna: organisms with a size of between 100 μm and 2 mm, for example, tardigrades, mites and springtails.
- Microfauna and Microflora: organisms with a size smaller than 100 μm, for example, yeasts, bacteria and fungi.

Most organisms in soil fall into the category of *microfauna* (more correctly called *microorganisms*). Microorganisms are characterized, by definition, by the fact that they only can be observed with a microscope, although there are exceptions. Soil microorganisms concern a large and diverse group of organisms, forming three distinct types, that is, *fungi*, *bacteria and archaea*. The majority of the soil organisms belong to the invertebrates, that is, organisms lacking a vertebral column.

According to another type of categorization, organisms are subdivided into three different domains. The first domain is the *Eukaryotes* (from the Greek words *karyon* (core) and *eu* (good)), that is, organisms with one or more cells with a cell structure including a cell core, internal membranes and cytoskeleton. Eukaryotes are subdivided into four kingdoms: animals (*Animalia*), plants (*Plantae*), fungi (*Fungi*), and protists (*Protista*). The second domain, the *Eubacteria*, comprises prokaryotes, organisms that do not have a cell core and whose DNA is available within the cell cytoplasm. The third domain consists of the *Archaea*, a group of single-celled prokaryotic organisms that distinguish themselves from the other domains by their extreme habitats and a biochemistry which is very different from the other domains. They live at high temperatures (sometimes over 100°C), in hot springs, in the deep sea, or in extreme alkaline, saline and acid water, and are therefore often called 'life's extremists'. Most of the *Archaea* live in anaerobic conditions.

13.2.2.2 Fungi

Fungi (from the Greek word *sphongos*, meaning sponge) are characterized by a chitinous cell wall, and often have a filamentous growth pattern. Unlike plants, they are not able to participate in photosynthesis (Lomako et al. 2004) and usually lack an efficient xylem- or phloem-like vascular system (Mihail and Bruhn 2005). Fungi are the largest organisms of the three groups of common soil microorganisms. Since the structure of fungi can range from individual cells to enormous chains of cells, there is a huge variation in size.

Fungi live for the most part in soil (e.g., yeasts and moulds) almost everywhere in the world, even in desert soils, and they play a very important role in most ecosystems. Around 70,000 fungal species have been formally described by taxonomists, but the true dimension of fungal diversity is still unknown and might be twice as big (Mueller and Schmit 2006). Many fungi have important symbiotic relationships, either mutualistic or antagonistic, with many other organisms (Perotto and Bonfante 1997).

The amount of fungal biomass in soil may be very high and one gram of soil may contain 10 to 1000 meters of hyphal threads.

13.2.2.3 Bacteria

Bacteria (from the Greek word *bacterion*, meaning 'small staff') are prokaryotic organisms (i.e., do not possess a cell nucleus) and differ from eukaryotic organisms (organisms with a cell nucleus), in many biochemical aspects. The most typical characteristics of bacteria is that they are unicellular, have no nucleus and no internal membranes. Usually they are one to a few micrometers in length (roughly about 50 times smaller than the width of a human hair) and have a wide range of shapes, ranging from spheres to rods to spirals. Bacteria are found everywhere, in every habitat, also in every soil in the world. They are the most numerous types of organisms present in soils. Most of the bacteria have not been characterized. Estimates of the number of bacteria in a gram of soil vary between 40 million (Whitman et al. 1998) to 1 billion (Conklin 2002). Bacteria are capable of very rapid reproduction, by dividing in two, when the conditions are favourable. One bacterium is capable of producing 16 million more in just 24 h (Bacteria Guide 2009); however, in soil the average growth rate of bacteria is in the order of two divisions per year resulting in maximum 4 cells from one.

Aerobic bacteria are most active in moist soils, but not in water-saturated soils, and at neutral soil pH, and where there is enough food available.

An important group of soil bacteria is formed by the *Actinomycetes.* They are fungi-like bacteria. They are smaller than fungi, and are also noted for a filamentous and branching growth pattern that results, in most forms, in an extensive colony, or mycelium, that look like spider webs. Mobility, when present, is conferred by flagella. Actinomycetes are perhaps best known for production of antibiotics. The number of actinomycetes might vary between 10 and 100 million per gram of soil.

Actinomycetes are responsible for the sweet, earthy smell of soil that some people dislike and others love. Some actinomycetes are important pathogens (i.e., a biological agent that causes a disease).

13.3 Organisms in the Groundwater

In the water-saturated groundwater zone, up to several hundreds of meters deep, organisms are present, mainly living under anaerobic conditions. Although the layman may be surprised to learn of the numbers of organisms in the groundwater zone, their abundance is not as numerous as in the upper soil. Most of these organisms are anaerobic bacteria (mainly using nitrate or sulphate as an electron acceptor) or *Archaea,* that is, methane producing organisms. In the aerobic groundwater, you may find also eukaryotic organisms including so-called *stygobionts* (from the Greek word *Styx*, the mythological river of the underworld) that are solely found in the subsoil aquatic environment. They are adapted to these specific conditions, for example, they are depigmented, have vermiform bodies and lack eyes. Their metabolism is reduced and the reproductive output is relatively small.

Compared to the organisms in the upper soil, not much is known about their habits, their interactions with abiotic factors, let alone their functions. It is generally accepted, however, that they contribute to many essential functions of the soil, for example, to the decomposition of organic contaminants an important process in the self purifying capacity of the soil, and important for the protection of groundwater quality.

13.4 Significance of the Soil Ecosystem

13.4.1 The Value of Soil Biology

With an abundance of more tan 10,000 different species and 1,000,000,000 (one billion) different organisms per gram of soil (Dykhuizen 1998; Torsvik et al. 1990), the abundance of bacteria is very high. This has led to discussions in the international research community on the necessity to protect soil Biodiversity. Redundancy is thought to be so high that the loss of some species might not result in function loss, although the relevance of microorganisms in soil is an issue which remains controversial in the international research community (Emmerling et al. 2002).

However, human life in a modern civilization, in its present form, would be impossible without a healthy soil ecosystem. Without a sound soil Ecological Health, the foundations of our social and economic systems would be undermined. From this perspective, the soil ecosystem has a tremendous economical significance. However, many of the services provided by ecosystems do not have an explicit value in our conventional market economy (Daily et al. 2000), and therefore environment and economy are still unequal partners (European Environment Agency 2008). But Ecological tests are used, such as ecosystem accounts, to describe the way ecosystems perform. However, it could be beneficial if some kind of monetary price could be attributed to Ecosystem Services. This would offer decision-makers the possibility to be able to weigh the value of Ecosystem Services alongside other social and economic information.

In 1997, the economic value of Biodiversity in general was calculated, resulting in the conclusion that it was two times the gross national product of the whole world (Costanza et al. 1997). In addition, in the Millennium Ecosystem Assessment very high values for Ecosystem Services were reported (UNEP 2005). The economic value of Ecosystem Services is still under discussion, but it is clear that it is very significant.

As mentioned in Section 13.1.3, the amazing performance of the soil ecosystem is largely unknown to the general public. Even with decision-makers dealing with contaminated sites, there is often confusion about the exact role of organisms in the soil, and the extent to which they fulfil their tasks. They generally realize, though, that human life is better off with a healthy soil ecosystem. Many decision-makers know about the functions that the soil ecosystem performs. However, not all are aware of the entirety of these functions and the powerful production for which these organisms are responsible.

13.4.2 Biodiversity

Biodiversity is the variety of life, in other words, the variety of the ecosystems, species and their genes. The definition of Biodiversity relates to the whole ecosystem of which the organisms are a part. In general, Biodiversity is approached from the perspective of its intrinsic value, that is, the significance of the entity by itself, independent of direct or indirect advantages for mankind. Typically, Biodiversity is assessed in terms of numbers of species. With the purpose to stress the importance of the concept, the United Nations declared 2010 to be *the International Year of Biodiversity*, giving space to celebrations in at least 80 countries, listed in UNEP Convention of Biological Diversity (2010).

It is generally acknowledged that Biodiversity has decreased during the last few decades. Therefore, the decline in Biodiversity is one of the eight threats mentioned in the EU Thematic Soil Strategy (Commission of the European Communities 2006). The European Soil Framework Directive, which is in the process of development, does not cover soil Biodiversity directly. Biodiversity will generally benefit from the measures proposed for other threats. This will contribute to achieving the objective of halting the decline in Biodiversity by 2010. Not enough is known about soil Biodiversity to give technical and political guidelines on how to handle and protect it. This is addressed in the Seventh EU Research Framework Programme, with a view to gaining a better understanding of the function of Biodiversity as an environmental service. This knowledge-building process will also be supported by ongoing initiatives under the Convention on Biological Diversity.

The urge to protect Biodiversity may be argued from different perspectives. For many fundamental biologists and ecologists, safeguarding Biodiversity is simply a normal aspect of dealing with our planet. They use the motto ‘live and let live’. Others support this idea from an ethical, spiritual or religious perspective. From an ethical or spiritual viewpoint, many humans are aware of the enormous pressure that the human population is exerting on the planet and the fact that we might experience repercussions if we do not improve our behaviour, including caring for all living organisms. Religious reasons may arise from the Christian-based principle that humanity has been entrusted by God with the planet’s stewardship, and that humans have the Christian duty to safeguard its precious balance. A rather extreme example of protection of the ecosystem from a religious perspective can be found among the Jains, followers of Jainism, the religion and philosophy originating in ancient India, who advocate the equality of all living organisms, no matter their appearance, size or function (Chapple 2006).

The significance of the preservation of Biodiversity was widely acknowledged on a global scale at the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro (often called the *Earth Summit*) (United Nations 1992). Cooperation among nations, in a spirit of global partnership, for the purpose of conserving, protecting and restoring the health and integrity of the Earth’s ecosystem, was the focus of the summit, a principle that still echoes in today’s environmental policy discussions. Besides the intrinsic value of Biodiversity as grounds for protection, the functional aspects of Biodiversity were also mentioned in the treaty.

Since the soil is inhabited by a large number of species differing widely in size, complexity and in their role in different processes, the different soil-dwelling species tend to be valued differently. It must be realized, however, that soil organisms live in a complex mutually dependent relationship. One organism serves as food for another organism, for example (see Fig. 13.2 for an example of a food web); intermediates in nutrient cycling performed by specific groups of organisms are essential for the survival of other species. In other words, different species need each other and the value of specific species must always be considered from the perspective of the functioning of the whole ecosystem.

There is a relationship between Biodiversity and the quality of Ecosystem Services (e.g., Griffiths et al. 2001). However, this relationship is very complex, and hard to quantify. Nevertheless, except for the sole protection of species, Biodiversity is an important criterion for the functioning of the soil ecosystem, or in other words, is an *indicator* of the quality of Ecosystem Services.

13.4.3 Ecosystem Services

13.4.3.1 The Significance of Ecosystem Services

As described in Section 13.1.2, the term *Ecosystem Services* refers to the benefits of ecological functions for the life of humans. Obviously, although the focus in this Section is on human benefits, organisms perform Ecosystem Services for their own well-being. They perform these tasks for feeding and, hence, for their energy supply, growth and reproduction. Generally speaking, huge numbers of organisms feed on large numbers of other organisms, their waste products, and the substances they release such as vitamins, amino acids and sugars.

In soil ecosystems, processes take place including the exchange of energy and chemical compounds between organisms and their environment. These processes lead to environmental circumstances that enable the existence of plants, bushes and trees. The overall Ecosystem Service related to these processes is the provision of agriculture and hence food, a prerequisite for human survival. Moreover, soil organisms are thus responsible for the existence of nature and green elements in urban areas. Organisms regulate the cycling of many important components. They are engaged in the element cycles and produce nutrients for plant growth in the carbon, nitrogen, and sulphur cycles, such as ammonia, nitrate, sulphate, and carbon dioxide. In aerobic soils, one important process is oxidation, and its main products are carbon dioxide and water. In anaerobic soils, microorganisms cannot use oxygen as an electron acceptor and instead use nitrate, sulphate or even organic contaminants, producing nitrogen, sulphide, carbon dioxide and methane. Although all these processes are performed by these organisms in their own interest, they are of invaluable importance to human life.

The significance of the processes performed by soil ecosystems can best be understood by investigating the performance of ‘natural soils’, that is, soils that have not been touched by human hands or machines, such as forest soils or prairie soils. Often these soils harbour a healthy ecosystem, since these soils are tilled and fertilized solely by the soil organisms.

In the Millennium Ecosystem Assessment (2005) a wide range of Ecosystem Services is listed, including provisioning services (e.g., the provision of medicines), regulating services (e.g., climate regulation), cultural services (e.g., recreation) and supporting services (e.g., soil formation). The value of several novel Ecosystem Services is under investigation, such as the ability of some soil fungi to produce iron chelators capable of extracting iron from crocidolite (blue asbestos), significantly reducing the potential to generate radicals and to damage DNA in the human body (Martino et al. 2004). Another example of an emerging Ecosystem Service focuses on the physical ground improvement methods with regard to the bearing power. Van Paassen (2009), for example, developed a novel ground improvement method, based on precipitation of calcium carbonate crystals (MICP) induced by the bacterial species Sporosarcina pasteurii. Since these crystals form sticking wedges between the sand grains, the strength and stiffness of the sand increases significantly.

An extended review of state of knowledge of soil Biodiversity, its functions, its contribution to Ecosystem Services and its relevance for the sustainability of human society is given in European Commission DG ENV (2010). In the present section, some 'basic' Ecosystem Services will be described in more detail.

13.4.3.2 Soil Structuring

A good soil structure facilitates root penetration, the exchange of gas (oxygen, carbon dioxide, methane and nitrous oxides) and water with the atmosphere, and the buffering of water. It also provides space for soil organisms. A well-structured soil is built up of aggregates, that is, individual soil particles that are joined into clusters. Such a soil is therefore relatively insensitive to crusting and clodding and, in the case of clay soils, to becoming sticky when wet. Well-aggregated soils are erosion resistant, since the aggregates are heavier than the individual soil particles. In dispersed soils the relatively light soil particles are sensitive to being blown away by wind or washed away by water. Moreover, the drainage is better in aggregated soils. Aggregates facilitate drainage, while having a high water storage capacity at the same time. Boyle et al. (1989) showed that untilled soil has more than three times the water infiltration capacity of mouldboard-ploughed soil. An aggregated soil facilitates easy seedling emergence and deep, prolific plant-root systems. Soil structuring indirectly influences heat and cold storage and, hence, even climate.

Many organisms contribute significantly to the loosening up of the soil. With regard to sustainable agricultural practices, reference is often made to the blessings of the soil ecosystem as a natural tillage performer. These blessings relate to good moisture conservation, high water infiltration, limited water run-off and the long-term building up of organic matter. Soils that are not burdened by the footsteps of humans and larger animals, traffic, agricultural machinery, etc., have a loose, slightly puffed-up structure.

Several glue-like substances such as gums and waxes, are produced by fungi and bacteria that are responsible for 'water-stable' aggregates, that is, aggregates that are stable when wet. Also the threads produced by fungi contribute to aggregate formation. Fungal hyphae, for example, physically bind soil particles together, creating stable aggregates that help increase water infiltration and soil water-holding capacity (Ingham 2000b; Van der Wal 2007).

Several soil-dwelling organisms are crucial for structuring the soil, for example, earthworms by digging tunnels (macro pores). Earthworm burrows enhance water infiltration. Earthworm tunnelling can result is an increase in water absorption by a factor of four to ten (Edwards and Bohlen 1996). Their tunnels, often lined with nutrient-rich organic compounds, can remain in place for years. Moreover, earthworms form clusters of soil material (*worm casts*), ranging in diameter from one or two millimetres to one centimetre, as the result of their excretions from the digestive tract after consumption of soil particles and microbes. A good population of earthworms can process 9,000 kg (20,000 pounds) of topsoil per year (Edwards and Burrows 1988). Earthworms also contribute considerably to the redistribution of metals from the deeper layers to the upper soils (Zorn et al. 2008).

To a lesser extent, the actions of smaller organisms such as enchytraeids and collembola are also important in soil structuring (Langmaack et al. 1999).

13.4.3.3 Humus Formation

Humus is a complex organic colloidal material with a dark colour. It is the end product of organic matter decomposition and originates from detritus, dead leaves and other plant materials, dead (micro)organisms and faecal material. It is very resistant to decay, sometimes lasting as long as 1000 years (Conklin 2002). Under natural conditions, humus is found on the soil surface and in the upper layers of the soil. It is intensively colonized by all kinds of organisms. Generally, old and young dead organic material mixes together on top of and in the soil, resulting in mixtures of different organic materials, partly resistant to decay and partly in the progress of decomposition.

Humus has excellent qualities with regard to the storage of moisture, and organic and inorganic compounds added to soil. It provides a habitat for plant roots and organisms. Therefore, it is extremely useful for plant growth and, hence, improves agriculture, horticulture, the development of nature, and the potential for plant and tree growth in urban settings. For this reasons it is often used as a soil conditioner and fertilizer, and also as a man-made humus called compost.

It is interesting to note that soil organisms contribute to humus formation, and thereby adapt their environment so that it functions optimally. An example of this phenomenon is the important role of earthworms in humus formation, given that the optimal habitat for these earthworms is, by the same token, a humus-rich soil layer. From this perspective, it can be concluded that, in addition to feeding relationships, one species or organism may profit from another species or organism.

13.4.3.4 Element Cycling and Nutrient Supply

All chemical elements are cycled within the biosphere and therefore are present in different forms. In the carbon cycle, gaseous carbon dioxide is incorporated in complex organic material by the process of photosynthesis in plants and algae. Plant material is used by organisms to feed upon, and after excretion the organic material is degraded by organisms into its elemental components and the whole cycle starts all over again. Plants use different elements for growth such as nitrogen (in the form of ammonium or nitrate), sulphur (mostly in the form of oxides), phosphate and other elements (nutrients). All the nutrients are intermediates in cycles: the nitrogen cycle, sulphur cycle, the phosphorus cycle, etc. (Stevenson and Cole 1999).

The cycles mentioned are performed by plants and organisms, and soil organisms especially play an important role. These cycles are tightly coupled. For example, when plants take carbon dioxide from the air, they are active within the carbon cycle. To make plant material, however, they need nitrogen and other nutrients to make the plant's components. Here, the plant components are intermediates in the different cycles; for example, proteins contain carbon, hydrogen, oxygen and nitrogen, and so proteins are intermediates in the four corresponding cycles.

Green plants and trees use carbon dioxide to produce plant material using CO_2 as a carbon source and light as an energy source (carbon fixation by photosynthesis). Plant material is used by soil organisms for growth and reproduction, resulting in carbon fixation in soils within the bodies of organisms and in the humus fraction of soil. The degradation of fresh organic material into carbon dioxide and of the humic fraction of the soil by soil organisms is an important part of the carbon cycle.

In the carbon cycle, organisms such as mites, springtails and earthworms mix and consume dead plant material left on top of the soil, and thereby redistribute the organic matter throughout the topsoil. Other organisms, such as potworms and nematodes consume plant litter further in the soil.

Thereafter, bacteria and fungi become the major players in the decomposition of organic matter in soils. This decomposition takes place in many stages. Fungi convert hard-to-digest and larger organic substances slowly into forms that other organisms can use (Ingham 2000b). They are able to degrade highly complex and resistant compounds such as cellulose, gums and lignins (Conklin 2002). Actinomycetes also help decompose complex organic matter (Conklin 2002; Sullivan 2004). Actinomycetes are abundant in soils in which the easily decomposed organic matter has already been decomposed and only the more resistant compounds remain. Actinomycetes and fungi are the most important decomposers of tough plant materials such as bark and woody stem material. Moreover, they are very effective in attacking tough, raw plant tissues such as cellulose and lignin.

Bacteria are heavily involved in nutrient cycling. Firstly, they are responsible for the production of ammonia and ammonium through the decomposition of organic material. Subsequently, other bacteria perform the process of nitrification, that is, the transformation of ammonium into nitrate. Nitrate is an important nitrogen source for plants. Under anaerobic conditions (in the deeper layers of the soil) denitrification (conversion of nitrate to molecular nitrogen) takes place. This process leads to the loss of available soil nitrogen and a subsequent loss in soil fertility. However, chemically bound nitrogen usually circulates many times between the soil and organisms before denitrification returns it to the atmosphere. Bacteria play important roles in the rapid consumption and degradation of easily degradable substances such as proteins, lipids and sugars, producing nutrients for plants and other organisms in the soil. Bacteria are also performers in the other nutrient cycles, making nutrients available to plants and other organisms by releasing sulphur, phosphorous and trace elements from organic matter or from soil minerals such as potassium, phosphorous, magnesium, calcium and iron (Sullivan 2004). Phosphate solubilizing bacteria (*Bacillus megateriumi*) (PSB) and earthworms (*Pheretima guillelmi* and *Eisenia fetida*) contribute to the phosphorus turnover and transformation in soil (Wan and Wong 2004).

Another important part of the nitrogen cycle concerns *nitrogen fixation*, that is, the conversion of nitrogen from the atmosphere into nitrogen-containing organic substances (Lindahl et al. 2007). This important process is performed by nitrogen-fixing bacteria in the soil and by bacteria that reside in the rhizosphere in symbiosis with plants of the *Leguminosae* (e.g., clover) family. These bacteria create nodules in leguminous plants in which they form colonies.

13.4.3.5 Cleaning Function

An important process in soil ecosystems is the degradation, and hence elimination, of organic contaminants such as pesticides and other anthropogenically introduced, potentially toxic organic contaminants, mainly through microbial degradation processes. This function provides natural cleaning of the upper soil layer and, hence, reduced organic contaminants load for the groundwater and surface water. From an anthropogenic point of view, this process is often referred to as the *cleaning function* (also called 'filtering function' or 'purification function') of soil, focusing on the provision of clean groundwater. This cleaning function is often sufficient to be able to use groundwater as drinking water without the need for any purification process. In other cases, mainly in urban areas, it facilitates the purification process for the waterworks.

Also in the groundwater are organisms that degrade organic contaminants. In in situ remediation of soils, the natural cleaning capacity is used explicitly. It is even stimulated by the addition of chemicals, enhancing the activities of the soil's microbial community, such as electron acceptors (nitrate, ozone), nutrients (nitrogen, sulphate), vitamins, or even specific organisms, in order to upgrade the soil microbial community.

The process of degradation in combination with dilution by transport processes is often called *Natural attenuation* (see Chapter 22 by Peter et al., this book, for details on Natural Attenuation).

It must be mentioned that the process time for degradation of persistent organic contaminants in soils might take a long time (half lives of more than half a year). Persistency is defined as a characteristic of a contaminant which pertains to the duration of its effectiveness. However, since these contaminants usually do not migrate fast through the soil, the time span between entering the soil and reaching the groundwater is often relatively long, which means that there is much time available for degradation. One drawback of the degradation of organic contaminants is that the degradation products sometimes result in the formation of compounds that are even more toxic than the mother contaminants.

Another process that supports the reduction of contaminant leaching into groundwater is the precipitation of heavy metals in the form of sulphides or phosphates, produced by anaerobic bacteria. Also the increase in pH, often resulting from biological activities in lower soil layers, results in a reduction of metal mobility and, hence, of leaching into the groundwater.

13.4.3.6 Disease Control

In most agricultural soils the incidence of plant diseases is a well-known problem. Soil organisms may play a role as the cause of, and the remedy against, crop diseases (phytophatology). Soil-borne pathogens belong to several different phyla such as bacteria, fungi or nematodes. They reside in the soil for brief or extended periods, and survive on plant residues or as resting organisms until root exudates reach them and allow them to grow. They may escape competition with other microorganisms by penetrating the roots. They either remain inside the plants until host death, or move outside the plants to infect other parts of the root or other roots (Haas and Défago 2005).

It is generally acknowledged that soil-borne diseases result from a reduction in the Biodiversity of soil organisms. Plant diseases can effectively be treated with chemicals, that is, fungicides, nematocides and bactericides. These chemicals, however, may have negative site effects such as affecting useful organisms and leaching into the groundwater.

By the time disease symptoms appear, disease pathogens are inside the plant and generally beyond control. Therefore, it is important to prevent the penetration of pathogens. Healthy soils are to a large extent able to prevent infection of plants, even though pathogens are present. These soils are called *natural suppressive soils*. The level of disease suppression is typically related to the level of total microbiological activity in a soil (e.g., Sullivan 2004).

Some fungi are able to trap harmful plant parasitic nematodes physically. The thick mycorrhizae network, for example, physically obstructs the penetration of root-feeding nematodes (Sullivan 2004). Furthermore, bacteria and fungi are crucial in natural disease control. Haas and Défago (2005), for example, demonstrated that some pseudomonades produce antifungal antibiotics, elicit induced systemic resistance in the host plant or interfere specifically with fungal pathogenicity factors during root colonization.

Fungi are not just active as pathogens. The first antibiotic isolated for human use was penicillin, obtained from the soil fungus *Penicillium chrysogenum*.

Beneficial organisms can be added directly to prevent plant diseases, but a common and sustainable procedure in plant disease control is also to stimulate the activities of microorganisms that are antagonistic to plant-parasitic nematodes by applying organic soil amendments to the soil (e.g., Akhtar and Malik 2000).

13.4.3.7 Energy-Related Ecosystem Services

It is a well known fact, also to the general public, that the presence of fossil energy sources such as coal, oil and gas in soil is the result of former biological life, often dating back several millions of years. From this perspective, the ‘supply of fossil energy sources’ is a special Ecosystem Service (if it can be called that), since the rate of fossil energy consumption by humans is orders of magnitudes higher than the production by the soil ecosystem. In other words: fossil energy supply could be considered as an Ecosystem Service provided in the past and available for a limited period of time.

Over the last few decades, the mitigation of the negative consequences of the use of fossil energy sources and the possibilities for using clean energy sources have made enormous gains in public interest. From this perspective, soil can make a contribution to the fixation of greenhouse gasses. A nice example is the fixation of CO_2 in soil organic matter, which is taken into account in the Kyoto climate treaty. A nice example of the application of soil organisms to avoid the methane produced in waste dump sites entering the atmosphere is by covering the dump with a layer of soil, so that the methane is oxidised into CO_2. Methane is a much more potent greenhouse gas than carbon dioxide, and reduction of its emission damps down climate change.

An extremely interesting development with regard to clean energy sources is related to the use of insoluble electron acceptors by anaerobic soil organisms. This is a natural process occurring in anaerobic soils in the degradation process of organic matter in the absence of nitrate or sulphate as electron acceptors. This process is being investigated now to produce so-called *Bio-electrochemical systems* (*BESs*) that use the electrical energy supplied by bacteria for energy production. In so doing, these bacteria, also so-called *Microbial Fuel Cells* (*MFCs*), produce electricity and can be considered as small batteries. These bacteria live on organic matter secreted from rice plants, and leave water and carbon dioxide behind. So far, plain electrical power production combined with wastewater treatment by Microbial Fuel Cells has been the primary application of BESs (Clauwaert et al. 2008; Stams et al. 2006). A wide variety of bacteria can participate in this electron transfer, and this phenomenon is far more widespread than previously thought (Rabaey et al. 2007).

13.4.4 Above-Ground Biology

Most of the Ecosystem Services described above (Sections 13.4.3.2, 13.4.3.3, 13.4.3.4, 13.4.3.5, and 13.4.3.6) take place beneath the soil's surface, but have a direct impact on developments above ground. This is most directly reflected by above-ground plant growth, either related to man-grown crops, as is the case in agriculture, horticulture or gardening, or related to herbs, bushes and trees in nature reserves, shoulders, and in fact on practically every bare patch of the planet. Indirectly, the presence of vegetation strongly impacts the presence of all other forms of wildlife, ranging from small insects on up to large animals such as deer or bears. The presence of trees and plants is an extremely important element in the daily life of almost every citizen, both in rural and urban areas. And for these trees and plants, a healthy soil ecosystem is an elemental boundary condition.

13.4.5 Agriculture

The relationship between agriculture and the soil ecosystem is a typical example of a love-hate relationship. Traditionally, farmers are very much aware of the crucial role that the subsurface biology plays in successful agricultural production. Although they might not fully comprehend the details of the magnificent biological and abiotic processes that are ongoing in soils, they generally are convinced of the necessity to care for the soil ecosystem. In fact, practically every function the soil ecosystem performs, described in Sections 13.4.3.2, 13.4.3.3, 13.4.3.4, 13.4.3.5, and 13.4.3.6, supports agriculture. From this perspective, the soil ecosystem could be considered as livestock.

However, since the 1950s, industrial agriculture, rooted in the massive use of technology and the application of chemical products, has spread worldwide. In this technological agriculture farmers did not dare to trust natural processes to enhance production, and therefore applied chemical and mechanical technology. Nutrients were applied, not by degradation of the organic residues of plants remaining from the previous growth season, but by the addition of fertilizer. Pest control was not carried out through a proper design of the farm, providing a place for pest-reducing organisms, but by the application of pesticides. Tension arose between fast-production imperatives and more sustainable ways of farming. Soil structure was no longer supplied by the activity of earthworms and other soil organisms, but was obtained by tillage, resulting directly in tilling the worms and indirectly by destroying their tunnels, drying out the soil, burying plant residues they feed on and exposing the soil to freezing. Tillage destroys the total natural habitat of soil organisms, leading to huge losses of organic matter, due to exposure of organic material to the surface and thereby to oxygen. In peaty soil especially, this can be seen very clearly.

Excessive nutrient loading and heavy machinery are also severe threats to the soil ecosystem. But, over the last decade, as more and more humans have grasped the impact of industrial agriculture and started questioning its long-term implications, interest in sustainable agricultural management systems has made a remarkable come-back. This has resulted in hundreds of books and publications on this topic, a rapidly developing market for organic products and many farmers shifting to more sustainable ways of growing food. In this sustainable agriculture, the application and use of ecosystem processes is very important and soil organisms have once more become the appreciated co-workers of farmers.

13.5 Ecological Risk Assessment

13.5.1 Principles

As outlined in Section 1.5.2, from a more general Risk Assessment perspective, risks have to do with *chance* and *effect*. Before describing Ecological Risk Assessment, it is useful to define the protection target, that is, the object which might be affected, in more detail. It is at this point that we reach the very core of Ecological Risk Assessment, namely, the definition of Ecological Health.

There is no universal concept of Ecosystem Health (De Leo and Levin 1997). Although there is a premise that natural systems are healthier than human-altered ecosystems (Lackey 2001), most ecological debates going on today focus on Ecosystem Health rather than on Ecosystem Integrity (Westra 1998). This is a sound approach since the unimpaired condition in which ecosystems do not show any influence from human activities, often seen as the basis of Ecological Integrity, is not necessarily the optimal condition for the soil ecosystem. Although many human activities are widely recognized as a major plague on ecological systems, humans may also, generally for their own benefit, enhance the soil ecosystem. From this perspective, the condition of pristine ecosystems may not always be considered as the ideal goal for a managed ecosystem, but it can be used as a reference point (or baseline, or benchmark) for an ecosystem.

Contrary to Human Health Risk Assessment, which focuses on one single (although extremely complex) species (i.e., human beings), Ecological Risk Assessment focuses on the total soil ecosystem including numerous completely different species (see Section 13.2.2). Protection of the soil ecosystem is important because of its intrinsic value (as was stated at the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro (United Nations 1992), but also because of the Ecosystem Services of the soil. Man depends on healthy soil for his survival and supply of food, energy and clean groundwater. For the provision of these services, organisms are important. Therefore, if we want to protect these services, we have to protect the organisms performing them. However, effects on the ecosystem cannot be assessed as the sum of the effects on all separate species involved. In fact, a major problem in Ecological Risk Assessment is the extrapolation of observations from individual and population levels to the ecosystem level (Eijsackers et al. 2008). Therefore, in an ideal Risk Assessment approach the ecosystem is considered from a *holistic perspective*, that is, as the whole system of mutually dependent organisms and abiotic parameters with regard to the whole series of Ecosystem Services.

Another complicating factor is that there are different stakeholders with regard to the soil ecosystem, who have different interests in the way the soil ecosystem functions. In other words, there is no general interest in, and not one ideal state for the soil ecosystem. This is another challenge for decision-makers in deciding which of the various preferences are to be adopted, see Rutgers (2008) who showed that different stakeholders in the Hoeksche Waard in the Netherlands have different preferences.

Many different Ecological Risk Assessment tools are used around the world for similar purposes. Therefore, Swartjes et al. (2008) made the case for harmonisation, that is, improving the international consistency in the technical parts of Ecological Risk Assessment procedures, at least at the European level.

13.5.2 Risk Characterisation

As explained in Section 1.5.3, from a more general Risk Assessment perspective, Risk Characterisation could be preceded by two steps, namely, the exposure assessment (representing ‘chance’ in Risk Assessment terms) and the Hazard Assessment (representing ‘effect’ in Risk Assessment terms). However, in Ecological Risk Assessment these steps are often combined. The reason for this is that there are no (or fewer) ethical obstructions, as opposed to Human Health Risk Assessment, to empirically investigating the relationship between the concentration in soil or in water and the effects on organisms or on Ecosystem Services. Moreover, the performance of soil ecosystems is less dependent on specific exposure scenarios than the performance of humans. The performance of ecological experiments also has the advantage over human health experiments that the effects are easier to measure. Small organisms show, for example, a lower activity or a specific percentage of fatality. Humans, on the contrary, show less concrete effects related to a specific organ, which will often only be revealed as sickness years after exposure to the contaminant.

Combined exposure and Hazard Assessment has the advantage that a direct determination of the relationship between soil concentration and ecotoxicological effects is much more reliable than a separate two-step determination. Such a separate determination would involve large uncertainties related to, firstly, the determination of a relationship between the soil concentration and exposure and, secondly, the determination of a relationship between exposure and ecotoxicological effects. As an example, the relationship between the concentration in soil and effects on earthworms could be considered. A simple experiment can be performed in which several earthworm populations are exposed to soils with an increasing contaminant level, and for each of which the activity of the different worm populations is measured. It is more complex, laborious and unreliable to measure the exposures (more precisely: the bioaccumulation) in the different earthworm populations and subsequently to determine the relationship between the bioaccumulated concentration and effects such as activity, on the earthworms. Moreover, for most small organisms in soils such as bacteria, for example, it is relatively difficult and expensive to measure the intracellular content of contaminants and, hence, to determine exposure.

In experiments, the exposure and hazard assessment are generally combined, since most often the effects are measured as endpoint. To this purpose, microcosms are often used, these are artificial, simplified ecosystems, which are used for simulating and predicting the behaviour of and effects on natural ecosystems under controlled conditions. An example is given by Ernst and Frey (2007), who performed a soil microcosm experiment, with the purpose to assess the uptake of mercury from various mercury-spiked food sources by two earthworm species.

It must be noted that for ethical reasons the harm to organisms, most certainly in the case of larger animals, must be weighed against the significance of the experiment for Ecological Risk Assessment. In the case of the use of larger test animals, alternative procedures must always be considered. In most countries, strict regulations apply to the performing of experiments on test animals (see Section 5.6.6.2).

13.5.3 Characteristics of Exposure

13.5.3.1 Oral and Dermal Exposure

Exposure models can be used to assess the exposure from contaminated sites on wildlife (see Section 13.7.5). As was mentioned in Section 13.5.2, however, the exposure itself is generally not determined in Ecological Risk Assessment, when considering soil organisms. For *insight* into Ecological Risk Assessment, however, it is useful to understand the exposure of the soil ecosystem. Typically, soil organisms reside in the (partly water-filled) soil pores in close contact with contaminants, practically all their lives. As a consequence, unlike human beings, soil organisms are more or less constantly exposed to contaminants. However, the intensity of exposure changes over time and is very different between species. Generally, two type of exposure occur, that is, oral and dermal exposure. Oral exposure happens through the consumption of other organisms, their excrements and soil particles with or without organic matter, which contain contaminants. This oral exposure is dependent on the gut conditions of the organisms (see, for example, Vijver (2005)). Dermal exposure mainly occurs via dissolved organic contaminants in the pore water. Some species either are exposed through dermal *or* oral exposure only, while other species are exposed through both oral *and* dermal exposure.

Ideally, the combined effect of all contaminants present on the ecosystem is assessed (e.g., De Zwart and Posthuma 2005). In the Netherlands, for example, the local toxic pressure of mixtures is included in the latest revision of the Dutch Soil Protection Act, by applying the ms (multi substance)-PAF method (Boekhold 2008), based on the Species Sensitivity Distribution concept (see Section 13.6.2).

13.5.3.2 Bioavailability

An important phenomenon with regard to exposure, most certainly in terms of dermal exposure, is bioavailability (Peijnenburg et al. 1997). Bioavailability relates to a specific fraction of contaminants in soil that is 'operational', that is, the *bioavailable* or *effective fraction*. This fraction is potentially available to be taken up by organisms, including plant roots, while other fractions are (temporarily) connected to solid soil phases such as organic matter particles, clay minerals, or (hydr)oxyde surfaces, and hence (temporarily) not available. The bioavailable fraction is mainly relevant with regard to dermal exposure. For oral exposure, the bioaccessibility of contaminants adsorbed to organic matter in the organisms controls internal exposure.

The bioavailability of a contaminant is a dynamic variable, dependent on the chemical characteristics of the contaminant and the soil, and on the characteristics of the organism involved. Bioavailability is related to pore water concentration and to the desorption or dissolution rate of a contaminant.

It is important to realize that the relevant bioavailable fraction depends very much on the time frame that is of interest. In fact, part of the contaminants are readily available, another part will become available within a period of days, weeks, months or years, while others virtually never will become available. The theory of bioavailability is strongly dependent on the type of contaminant. A very useful classification with regard to bioavailability characteristics relates to the two main classes of contaminants, that is, metals versus organic contaminants.

There is, however, some political controversy with regard to bioavailability. This controversy relates to the time factor. On the one hand, a higher bioavailability of metals will do more harm to the soil ecosystem. On the other hand, since a higher bioavailability usually (but not always) coincides with a higher potential for leaching, the contaminants will disappear from the ecologically most important soil upper layer in a relative short timeframe, after which the ecosystem can recover. Using the same reasoning, immobile contaminants are less bioavailable, but will reside longer in the ecologically most important soil upper layer. As long as bioavailability is low, the effects may be limited, but when environmental conditions change, for example, by changing the land use, they may become available again and show their effects then. It is a political dilemma grounded in what is worse: a big adverse ecological impact for a short period, a smaller adverse ecological impact for a longer period, or an unknown ecological effect later on. This question is even more difficult to answer when other protection targets (the groundwater that is impacted by leaching; human health that is impacted by contaminants in vegetables) are also taken into consideration. Given the diversity of soil ecosystems and the complex relationship with the many other (abiotic) factors in soil, there is, of course, no uniform answer to this question. The main message here is that the risk assessor and risk manager must, for the sake of sustainability, generally focus on long-term risks.

The bioavailability of contaminants is closely related to the soil properties. The absorption characteristics of the soil are, in addition to the bioavailability of harmful contaminants, also important for nutrient supply and the bioavailability of essential metals. Therefore, soil characteristics such as organic matter content and type, clay content, the content of manganese, aluminium and iron (hydr)oxides are important in buffering both useful chemicals as well as contaminants. They strongly influence sorption and, hence, bioavailability. The type of organic matter is important for at least two reasons. Firstly, organic matter as part of the solid phase reduces the groundwater concentration and, hence, bioavailability, while adsorption to *dissolved organic matter* increases the total concentration in pore water and may increase bioavailability. Second, different types of organic matter have different sorption affinities. Cuypers (2001), for example, showed that the bioavailability was much higher in the amorphous dissolved-organic-matter domain than in the condensed dissolved-organic-matter domain.

With regard to ecological protection, pH has special status. The presence of a high concentration of hydrogen atoms has huge impact on metals and metalloids adsorbed in and on all sorbents. Therefore, pH also controls the bioavailability of both useful chemicals and contaminants in soils. As a consequence, the acid buffering capacity of soil is an important factor for a good soil quality.

As a consequence, the absolute concentration does not play an important role in Ecological Risk Assessment. However, the ‘bioavailable concentration’, as a measure of exposure, is much more important with regard to normalisation of effect concentrations in soil (for example a Soil Quality Standard; a concentration that marks the difference between ‘no risk’ and ‘possible risk’). Imagine, for example, that an ecologically based Soil Quality Standard for a specific contaminant is 100 $mg/kg_{dry\ weight}$, and that the actual bioavailability at a specific site results in an actual exposure that is 50% lower than the prevailing bioavailability related the experimental conditions under which the Soil Quality Standard is set as 100 $mg/kg_{dry\ weight}$. When a linear relationship between ecological impact and concentration in the bioavailable fraction is assumed, it seems fair to increase the acceptable concentration for this specific site by a factor of two, that is, to compare measured concentrations with an ecologically based Soil Quality Standard of 200 $\mathrm{mg/kg}_{\text{dry weight}}$.

In spite of decades of research, however, the concrete application of bioavailability assessment procedures, certainly with regard to Soil Quality Standards, has so far been limited. The reason for this is that bioavailability is a highly site-specific process. Several countries have implemented practical guidelines for correcting Soil Quality Standards for bioavailability (Carlon and Swartjes 2007). In the Netherlands, for example, *soil type correction relationships* are used for this purpose on the basis of the organic matter content (organic contaminants), and the organic matter and clay contents (metals) (Ministry of VROM 2008). Although the use of these relationships for the bioavailability correction of Soil Quality Standards lacks a strong scientific foundation, it is believed that these corrections are better than totally ignoring bioavailability. Further research is on its way concerning the application of different mechanistic models to tackle the problem (e.g., Koster et al. 2006).

In Hodson et al. (Chapter 16 of this book), the principles of bioavailability are described in detail.

13.5.4 Endpoints

One important aspect in Ecological Risk Assessment is the definition of criteria for assessing the ecological state of the soil (*Soil Ecological Health*). For this purpose, relevant *endpoints* for testing must be selected, generally a protection target at a lower trophic level than the soil ecosystem. These endpoints play a role in the empirical derivation of the relationship between soil concentrations and Ecological Health (or with ecological effects) such as, for example, for the derivation of Species Sensitivity Distributions (SSDs), see Chapter 14 by Posthuma and Suter, this book. Effects are often related to, in ascending order of gravity, no-effect levels (No Observed Effect Concentration, NOEC), lowest observed effects (Lowest Observed Effect Concentration, LOEC), all kind of effects (Effect Concentration, EC), hazardous effects (Hazardous Concentration, HC), or death (Lethal Concentration, LC). Decision-makers may select an appropriate protection level in terms of a specific percentage that apply to one of the endpoints, for example, Hazardous Concentration for x% of the organisms (HCx) at which x% of the organisms are hampered (Posthuma et al. 2002).

The selection of endpoints is also an important issue in site-specific Ecological Risk Assessment. In these applications, a predescribed acceptable level of protection of the appropriate endpoint must be defined. Since the range in which the physical and chemical factors vary is strongly dependent on climate, soil type, soil conditions and land use, it seems fair that the requirements for specific soil ecosystems should differ by countries and regions. Often, endpoints are related to a single species such as activity, reproduction rate, and mortality related to one type of organism. It is also generally recognized that a healthy microbial population in soil is an indication that the soil is in good condition. This means that one relevant endpoint could be ‘the presence of a large, active, diverse microbial population’ (e.g., Bloem and Breure 2003). More specific endpoints for the effects due to cadmium uptake are used by An (2004), who measured the seed germination and seedling growth (shoot and root) for sweet corn, *Zea may*, wheat, *Triticum aestivum*, cucumber, *Cucumis sativus*, and sorghum, *Sorghum bicolor* grown on cadmium-amended soils.

A popular endpoint is the number of species, since it is generally accepted that there is a positive relationship between Biodiversity and Soil Health (Tilman et al. 1996, Griffiths et al. 2000). In addition, the effect on ecological processes (and therewith Ecosystem Services) is recognized as a meaningful ecological endpoint. These endpoints are mostly assessed with regard to the percentage of species or processes (Ecosystem Services) affected ($\text{PAF}_{\text{species}}$, Potentially Affected Fraction of species; or the $\text{PAF}_{\text{processes}}$, Potentially Affected Fraction of processes, respectively) (Posthuma et al. 2002), or via the inverse criterion, the percentage of species or Ecosystem processes (services) protected (that is, *not* affected).

At the more detailed level, however, the criterion ‘number of species’ or ‘effects on Ecosystem processes (services)’ needs specification with regard to the term ‘affected’. With regard to the number of species, different possibilities for ‘affected’ exist, that is, in ascending order of gravity: reduced activity (although large numbers of diverse species of microorganisms can survive in soil in an inactive or resting state (Conklin 2002), diminished reproduction, and death. With regard to Ecosystem Services, a range of different type of effects can also be distinguished.

An integral approach to ecosystem quality and the effects of stresses on the ecosystem starts with a proper description of the system. One approach is to focus on the food web: who eats whom, expressed in the flux of energy and nutrients between different groups of organisms, used by Hunt et al. (1987). This approach requires a detailed data set and, consequently intensive monitoring, with repeated measurements taken over the course of a year.

Another approach is to derive stable species compositions in ecosystems by using the allometric approach. Here, the degree of stability is calculated from the numbers and the mass of individual organisms, and thus from the total biomass of species and groups of species. The relationship between the place of a species in the web, and its abundance in numbers and mass, indicates the stability of the web. The approach has been applied in terrestrial ecosystems (Mulder 2006; Mulder et al. 2005).

Another approach is to derive criteria for the stability and efficiency of an ecosystem. A system is unstable when insufficient prey is available to feed the predators and vice versa. A system is inefficient when there is too much prey available, because then the available prey is not completely converted. This may lead to the formation of reservoirs of compounds such as storage of organic material because of the absence of predatory or degrading organisms.

Properly defined, stability and efficiency criteria for ecosystems and food webs enable the description of a ‘good ecological status’ (EU-WFD) and ‘sustainable use of ecosystem processes’ (EU-SFD) in an operational way (Breure et al. 2008).

However, endpoints generally on the same weight for every species and Ecosystem Service, and do not include information about the organisation of an ecosystem. It does not, for example, account for the impact on species that are very important for a whole chain of other species. Otherwise, species in soil might be redundant (e.g., Buckley and Schmidt 2001). The many processes (Ecosystem Services) are especially strongly interlinked. A healthy ecosystem has a complex organisation, with a less dominant role for opportunistic species. From this perspective, it is complicated to define relevant endpoints. The performance of relevant *indicator species*, however, could serve as an endpoint.

13.5.5 Other Stress Factors

13.5.5.1 Ecological Impact

It must be noted that the vitality of the soil ecosystem is not only impacted by soil contamination, but also depends on other environmental factors. Some examples are an unfavourable pH regime (usually a low pH, acidification), drought, soil compression, sealing, soil organic matter loss and, last but not least, human presence.

At any contaminated site, the risk assessor must be aware of these other stress factors for at least two reasons. Firstly, when ecological effects are observed, not all of the effects can be attributed to exposure to contaminants in soil. Secondly, in most cases it is not effective to reduce the contaminant load or bioavailability by means of Risk Management, when other stress factors are dominating ecological effects. This is, however, a policy decision.

A multivariate approach can be applied to filter the effects of contaminants from the many interacting factors at the ecosystem level (e.g., Van Straalen and Van Gestel (2008), with the purpose of assessing ecological risks for diffusively contaminated sites in the Netherlands).

13.5.5.2 Soil Type, Properties and Structure

Independent of soil contamination, soil life depends on soil type and the prevailing conditions in the soil. First of all, soil structure must allow for residence possibilities for organisms (the habitat function). Although many organisms create their own holes, cracks and tunnels, one soil is better suited for this purpose than another soil, depending on particle size and format (and hence pore structure) and aggregate building properties. It is interesting to note that the activities of one group of organisms create physical living conditions for other organisms, for example, earthworms dig tunnels which facilitate root penetration in soil.

Many man-made soils (*Technosols*) provide poor living conditions for organisms because of the presence of soil-foreign materials such as debris and tar. These materials create physical and chemical obstacles for organisms (e.g., Rutgers (2008), who showed several ecological impacts on the soil ecosystem in the Demmerikse polder in the Netherlands, an area with an anthropogenic layer of soil on top of peat soil, including soil-foreign materials).

Water management of the soil is an important factor. Most organisms need water to survive. However, water saturation in soil pores hampers gas exchange (oxygen supply). This has a negative impact on the majority of the organisms in the soil, which is aerobic and needs to exchange gas, that is, oxygen and carbon dioxide, with the atmosphere. As a consequence, both soil structure and water management are important factors for gas exchange processes and thereby for the proper performance of the habitat function.

Tillage may improve the soil conditions for ecological activities, for example, by stimulating oxygen and water supply. However, tillage also has a negative impact on the soil ecosystem, for example, by physically damaging the tunnels of larger soil-dwelling organisms such as earthworms and these organisms themselves. Sláviková and Vadkertiová (2003), for example, showed that the number of yeasts (from the fungi kingdom) was about ten times higher in forest soils than in tilled agricultural soils. For this reason non-till systems are increasingly gaining in application.

In conclusion, a healthy soil that stimulates vitality has a good combination of structure building possibilities, availability of organic material and other food sources, gas exchange and water management.

13.5.5.3 Food Supply

Since organisms need food, one extremely important factor is the availability of organic material. The total soil food web is based on the degradation of detritus, dead organic material which serves as food for soil organisms. The easily degradable organic fraction will be degraded by higher organisms and bacteria; the more recalcitrant compounds, including the humus fraction, is slowly degraded, after an initial attack by fungi that are able to convert such material to useable food for other organisms. Therefore, both the organic matter content of soil and its quality are important factors in soil vitality. Detritivores consume dead organic material, while phytophages feed on living plant roots and plant material. Since predators feed on other living organisms, they depend on the success of other organisms. Besides carbon, organisms need nutrients such as nitrogen, phosphate and sulphur which are recycled in the element cycles, coupled with the degradation of organic material.

13.5.5.4 Sealing and Compaction

In many parts of the world there is a significant increase in urban areas, in which sealed surface dominate. Sealing, that is, covering the surface of the soil with asphalt, concrete, paving stones, and also buildings and greenhouses, generally has a substantial impact on the soil ecosystem. The adverse effects differ from species to species and range from negligible to disastrous. Since the biggest problem caused by sealing is the blocking of the water supply and, depending on the degree of permeability (possible joints, cracks), on gas exchange, the ecological effects depend on the closed sealed surface area in relation to the presence of cracks and unsealed areas. The effect of sealing on Ecosystem Services is that the processes to be performed by organisms will be hampered by the drying out of the soil, which makes it an unfavourable habitat. If the oxygen supply is also insufficient, this might lead to anaerobic conditions that change the processes completely because of a changing community structure.

From an ecological perspective, an important boundary condition for spatial planning should simply read: as much bare or green-covered areas as possible (or as few sealed surfaces as possible). Moreover, although this is much more relevant for macro fauna, green surfaces must be as fully connected as possible, since there is a relationship between surface area and Biodiversity.

Compaction of soil, due to heavy agricultural equipment or intensive human presence, results in a reduced pore fraction of the soil. This impacts the soil ecosystem in different adverse ways. First, the organisms lose part of their habitat. Second, the exchange of gas and groundwater is hampered.

In areas where sealed surfaces alternate with green surfaces (see Fig. 13.3 for an aerial view of the Curium-LUMC Academic Centre for Child and Youth Psychiatry, Oestgeest, the Netherlands and surroundings) it is difficult to assess the impact on the soil ecosystem, for example compared to the same situation without any development on the site. On the one hand the soil ecosystem may be seriously hampered by both sealing and compaction. On the other hand, the spatially distributed green surfaces provide opportunities for ecological life in soil and allow transport of water and oxygen, partly also in lateral direction, so that ecological activity will thrive to some extent.

Fig. 13.3 An aerial view of the Curium-LUMC Academic Centre for Child and Youth Psychiatry, Oestgeest, the Netherlands and surroundings, as an example of an area where the soil ecosystem is hampered through sealing and compaction, but where soil ecological activity will thrive to some extent because of the spatially distributed green surfaces (Source: Curium-LUMC; reproduced with permission)

13.5.6 Political Awareness

Obviously, the significance of the soil ecosystem as a protection target with regard to contaminated sites is much more difficult to comprehend than the significance of human health or of groundwater quality. However, both protection and recovery of the soil ecosystem are crucial from the two different perspectives described above, that is, Biodiversity (Section 13.4.2) and Ecosystem Services (Section 13.4.3). The combination of the important role of soil life and the public lack of appreciation of this make intensive communication a priority in order to justify ecological protection and, for the political agenda, the costs involved for the tax payer. Insight into the important functions that are crucial for society, that is, the Ecosystem Services, and into the amazingly complex way in which organisms perform these functions, might lead to promotion of ecological protection as a political priority in several countries that do not yet include this protection target in Risk Assessment and Risk Management.

Recent developments in soil and water policy, substantiated in the EU Water Framework Directive (EC 2000) and a preliminary Soil Framework Directive (EC 2006), are aimed at a more integral system approach. This approach targets 'Good Ecological Status' for water and 'sustainable use of ecological processes' for soil, in short, stable, optimally developed and well-functioning ecosystems. Data for the desired ecological quality need:

- to be derived for site-specific situations;
- to be integrated with generalized aims using broadly accepted criteria;
- to enable integral assessment of the total environmental stress on the ecosystems.

The initial descriptions of ecological quality, based on species compositions, were made for aquatic ecosystems: the 'RIVPACS approach' (Wright 2000). This approach compared actual and optimal species compositions, but did not take into account the interactions between different species, and between species and abiotic components of the ecosystem. A similar approach has been proposed and developed as SIVPACS (SOILPACS), its terrestrial counterpart (Spurgeon et al. 1996). This approach has now been backed up, in the Netherlands, by data obtained from the ecological monitoring of soil (Rutgers et al. 2008).

All plants, animals and microbes, and all soil, water and air within an ecosystem, interact with each other. This results in combined abiotic and biotic processes (life-support functions, LSFs), comprising element cycling (mainly carbon and nitrogen), production of oxygen, preservation of clean water, and degradation of organic matter and organic contaminants. Governments increasingly realize that society depends on properly functioning ecosystem processes (Ecosystem Services) in order to survive. They are concerned with obtaining a scientific foundation for legislation that is based on the assessment and ranking of different types of environmental stress and ecosystem functioning (Beck et al. 2005).

With regard to the position of ecological protection on the political agenda, communication between the scientific community (e.g., the eco(toxico)logists) and decision-makers is highly challenging. In many countries, intensive communication between decision-makers and scientists has taken place, often resulting in the implementation of national regulations. But since the extent of the benefits of Ecosystem Services and, even more so, the way these services are to be maintained, are so complex, this communication has not always been effective. In many cases this has resulted in non-transparent ecological protection levels. In addition, the remediation costs of contaminated sites in order to regain an ecologically acceptable level, that is, reconstitution of a healthy subsoil, could be one reason that decision-makers are reluctant about focussing on ecological protection. Moreover, lack of knowledge in terms of ecological protection levels has sometimes resulted in over-conservatism in setting ecological protection levels. It is claimed that ecotoxicologists masquerade value-based assumptions as science (Lancaster 2000). Moreover, scientists are accused of incorporating beliefs, morals, values and ethics as properties of ecological systems (Kapustka and Landis 1998). Indeed, the difficulty of communication may well enhance the subjective opinions of scientists. Many experts, in turn, are unhappy about the fact that, in spite of many years of collaboration, their decision-makers have not been able to comprehend the blessings of soil ecology to a full extent and, hence, have been unable to respond with appropriate political measures.

In conclusion, communication between decision-makers and scientists is essential for appreciating the protection and recovery of the soil ecosystem and setting appropriate protection levels. Choosing the soil ecosystem as a protection target and the determination of an appropriate ecological protection level strongly impact Risk Assessment procedures such as the value of ecologically-based Soil and Groundwater Quality Standards. What is often overlooked is that Risk Assessment procedures are much more sensitive to these policy choices than to many scientific issues. Therefore, communication between scientists with a talent for communicating about practical implementation and decision-makers with scientific understanding is extremely important for the appropriate positioning of ecological protection and recovery on the political agenda and for the determination of appropriate ecological protection levels.

In Europe (Carlon and Swartjes 2007; Römbke et al. 2005) and in the United States, Canada and Australia (Barron and Wharton 2005), there is a trend that shows an increase in political interest in ecological protection with regard to the management of contaminated sites. In the Netherlands, the protection of the ecosystem was explicitly included in the definitions of soil ('. . . including . . . and organisms') and soil protection ('. . . that imply a reduction of the threat to the functional properties that the soil has for . . ., plant and animals'). Soil Quality Standards derived from SSDs were proposed by Van Straalen and Denneman (1989). In 1994, the Dutch Soil Protection Act was extended with the addition of a procedure to determine the urgency of remediation, among other things, on the basis of site-specific risks for the soil ecosystem (Swartjes 1999). This procedure was significantly improved by the use of the TRIAD approach (see Chapter 15 by Rutgers and Jensen, this book). In the German Federal Soil Protection Act of 1998, reference is made to 'harmful soil changes'. In the UK, a tiered approach for assessing the risks to the soil ecosystem has been proposed, in which conclusions have been drawn on the basis of the 'weight of evidence' approach (Weeks and Comber 2005). This includes the impact on ecological soil functions (Römbke et al. 2005). In Europe, 11 countries consider the protection of ecological receptors, although only in a few of them formal ecological Soil Quality Standards have been adopted (Carlon and Swartjes 2007).

In Canada, federal ecological Soil Quality Standards for 87 contaminants were launched in 1999 (Canadian Council of Ministers of the Environment 1999). Australia established ecological Soil Quality Standards in 2003 (National Environmental Protection Council 2003).

13.6 Ecological Risk Assessment in Practice

13.6.1 Soil Quality Assessment

Ecological Risk Assessment is an important process in dealing with contaminated sites. Like all other Risk Assessment tools, Ecological Risk Assessment tools can be used for Risk Characterisation (risk appraisal for existing contaminated sites; see Section 1.5.3) and for Risk Management (recovery of soil quality). Although Ecological Risk Assessment tools do not have a long history when compared to Human Health Risk Assessment tools, there are many sophisticated Ecological Risk Assessment tools available today, worldwide (see, e.g., Carlon and Swartjes (2007) for an overview of ecological risk tools in the European Union). Many of these tools have been intensively used, tested and validated. Ecological Risk Assessment tools share several characteristics with Human Health Risk Assessment tools, for example:

- They have a limited reliability, since soil ecology is characterized by a huge temporal and spatial variability and the quantitative extent of many ecological processes are largely unknown.
- Worldwide, many different tools exist for similar purposes.

In spite of the limited reliability, Ecological Risk Assessment is an extremely useful process for understanding the threats to the soil ecosystem from soil contamination and supporting decisions on contaminated sites. However, a meaningful Risk Assessment requires that risk assessors be aware of the limitations. They must strive after a smart balance between relatively unreliable outcomes, and concrete and practical conclusions. Therefore, risk assessors have an important responsibility in communicating the outcomes of Ecological Risk Assessments, both in terms of scientific meaning and as restrictions.

So far, Ecological Risk Assessment has been based on statistical techniques that do not take into account the interactions between organisms and the relative importance of specific organisms for the realisation of soil processes. In the majority of the present models, organisms do not interact and they all have the same ecological importance.

It is not easy to incorporate ecological insight into routine Risk Assessment, but research on that subject is on its way. This research is now focussing on food web modelling based on monitoring data, which in the future should lead to Risk Assessment tools based on these ecological insights (e.g., Breure et al. 2008; Mulder et al. 2005).

A way to account with uncertainties in Ecological Risk Assessment is following a probabilistic approach. An example is given in Jager et al. (2001), who calculated probability density distributions of the PEC/PNEC (Predicted Effect Concentration, representing exposure divided by Predicted Non-Effect Concentration, representing effects) ratios for dibutylphthalate. The authors concluded that the deterministic PEC/PNEC ratios are worst case, generally higher than the 95th percentile. The disadvantage of a probabilistic approach is that a policy choice needs to be made for the level of acceptability, in terms of a specific percentile of the probability density function of the PEC/PNEC ratio. Although this offers a more sophisticated way of dealing with acceptable risks, there are no objective criteria to underpin this choice.

13.6.2 Soil Quality Standards

Statistical relationships between the soil concentration of contaminants and the fraction of species potentially affected in a contaminated soil are available (*Species Sensitivity Distributions; SSDs*). Analogously, relationships between soil concentration and the fraction of Ecosystem Services potentially affected in a contaminated soil are available (*Functions Sensitivity Distributions; FSDs*). Based on these relationships, soil concentrations can be derived which relate to a specified percentage of the fraction that is (potentially) affected and which can be selected as Soil Quality Standards.

It is the task of decision-makers to define the protection level, that is, the percentage of the affected fraction that is acceptable. However, since it is difficult to arrive at any insight into the consequences of the loss of a specific fraction of the organisms, or Ecosystem Services, this decision must be made in close cooperation with eco(toxico)logical experts. Examples of ecologically-based Soil Quality Standards are given in Verbruggen et al. (2001) for the Netherlands and in Fishwick (2004) for the UK. In the United States, there are several bodies that derived ecological Soil Quality Standards within a time span of seven years, doing this around the year 2000 (Barron and Wharton 2005), for example:

- the Oakridge National Laboratory (Efroymson et al. 1997; 54 contaminants);
- the New York State Department of Environmental Conservatism (NYSDEC 1998; three contaminants);
- the USEPA Region 5 (US Environmental Protection Agency (1999); 215 contaminants);
- the Oregon Department of Environmental Quality (Oregon Department of Environmental Quality (2001); 137 contaminants);

- the USEPA Region 4 (US Environmental Protection Agency (2001); 132 contaminants);
- the Texas Natural Resource Conservation Commission (Texas Natural Resource Conservation Commission (2001); 63 contaminants);
- the USEPA Office of Emergency and Remedial Response (US Environmental Protection Agency (2003); nine contaminants).

In Posthuma and Suter (Chapter 14 of this book) a detailed overview is given of the position of Species Sensitivity Distributions (SSDs) and Functions Sensitivity Distributions (FSDs) with regard to the derivation and use of Soil Quality Standards.

13.6.3 Site-Specific Risk Assessment

Site-specific Ecological Risk Assessment has undergone extensive development in a few countries in Europe, in the USA and Canada, ever since it was targeted as an important gap in Ecological Risk Assessment at the end of the 1990s. Today, a few promising methodologies exist that can rely on a few years of using and testing. One of the pitfalls, however, is that many site-specific ecological procedures focus on a single species (the single-species approach).

In the most appropriate and efficient approach for site-specific Ecological Risk Assessment, these tools are used in a tiered approach as for any other protection target. In a tiered approach the assessment becomes more site-specific and less conservative, and hence more complex, tier by tier. The motto is: simple when possible and complex when needed. An elegant procedure concerns the so-called TRIAD approach (Jensen and Mesman 2006). In this approach, the status of three different elements that relate to ecological damage, namely, of the soil chemistry, toxicology and ecology, is investigated (*multiple weight of evidence*). In Rutgers and Jensen (Chapter 15 of this book) the TRIAD approach is described in detail. Moreover, Chapter 15 includes a detailed overview of the theory and procedures for site-specific Ecological Risk Assessment, in general.

13.7 A Closer Look into Ecological Risk Assessment

13.7.1 Resilience and Recovery

Recovery often refers to the ability to return to a previous condition, that is, the condition before exposure to contaminants in the soil took place, when this condition has weakened or at least seriously has been deteriorated. *Resilience* can be defined as the speed of return to this previous condition or as the magnitude of disturbance that the ecosystem can absorb. Resilience and recovery are often, but not always, related to each other. Since both capabilities support the functioning of the ecosystem as a whole, including the Ecosystem Services and, hence, indirectly support human life, resilience and recovery are sometimes described as Ecosystem Services in and of themselves. In general, both qualifications are used for the ecosystem as a whole, since the huge diversity of organisms offers different kinds of possibilities for resilience and recovery. However, it also can be applied to a specific species or Ecosystem Service. It should be noted that since the mobility of soil organisms is much less than that of aquatic organisms, recovery times in soil are generally much longer than in an aquatic environment.

One important aspect of Ecological Risk Assessment, although difficult to assess, is the timeframe during which an ecosystem can be impacted by a specific soil contamination, while resilience and recovery are working to restore the ecosystem (Eijsackers 2004). In other words: up to what level of contamination and for what time span of exposure is ecological damage reversible?

Another important factor in Ecological Risk Assessment is the time span needed for returning to the previous position. Periods for restoration differ to a large extent among species. In general, the time spans for total recovery of vegetation and most microorganisms may require years to decades, while for larger animals this might take decades to centuries.

13.7.2 Adaptation

Adaptation is the ability of a community of organisms to improve its resistance to a threat such as soil contamination. Adaptation can either relate to the behaviour of the organisms (reducing exposure) or to the genetically controlled higher resistance of specific organisms (reduced effects or no effects). Adaptation can be regarded on an individual scale or from the perspective of a group of organisms. Individual organisms, specifically larger organisms, may learn to deal with soil contamination, but it is not expected that this significantly mitigates the corresponding adverse effects.

Adaptation from a population perspective is much more promising. It is well known from the medical use of antibiotics, for example, that there is a part of the bacterial population that is insensitive, and will survive and produce descendants that form a largely insensitive population. Madsen et al. (1992), for example, showed that at a forested site in the northeastern United States, metabolic adaptation to the presence of PAHs was evident; radio labeled naphthalene and phenanthrene were converted to $^{14}CO_2$ in core material from inside but not outside a plume of groundwater contamination.

Pollution-Induced Community Tolerance (PICT) is a tool to assess the adaptation capability. The concept of PICT covers the issue of causality better than classical ecological community response parameters such as species densities or species diversity indices (Boivin et al. 2002). Moreover, PICT relates to a more relevant level of ecological organization, namely, the community.

There is not much known about the power of ecosystems to create ‘inert populations’, not to mention ‘an inert ecosystem’, that can stand up against exposure to contaminants in soil or groundwater. Posthuma et al. (1992) found that the cadmium excretion efficiency of *Orchesella cincta* (L.) (Collembola) populations were significantly more efficient for organisms that were exposed to higher concentrations for a longer period, probably due to genetic differences. Sokhn et al. (2001) showed, in spite of a reduced microbial activity and an incomplete mineralization of phenanthrene, the presence of phenanthrene-degrading, copper-resistant and/or -tolerant microorganisms, at copper levels in between 700 and 7000 mg/kg_{dw}, indicating the presence of phenanthrene-degrading, copper-resistant and/or -tolerant microbes. In other words, adaptation does take place, but the system might become less efficient: Adaptation has a cost. This cost is related to energy requirements to cope with contaminants, for example, production of extra stress proteins, or energy to pump out a contaminant. Other microbiologist made the same observations earlier, for example, Doelman et al. (1994).

Adaptation is usually incorporated in generic Risk Assessment through the derivation of ecological generic Soil and Groundwater Quality Standards. This is done by following the *added risk approach* (Struijs et al. 1997), that is, assuming that the natural background concentration does not impose risks, since the ecosystem is adapted to these long prevailing conditions.

For site-specific Ecological Risk Assessments, risk assessors must account for possible adaptation. Biological *measurements* implicitly include the influence of adaptation. But critical soil concentrations, even when normalized for their actual soil properties, do often not apply for any adaptation.

13.7.3 Land Use

Unlike Human Health Risk Assessment, *land use* does not always significantly impact exposure of the soil ecosystem. However, one should not overlook the impact of subsurface building activities, infrastructure, agriculture and industrial areas on Ecological Health. The main difference with Human Health Risk Assessment is that human behaviour and residence time strongly influence the exposure, while the soil ecosystem is more or less constantly exposed to contaminants in soil, independent of the land use. One exception is the land use ‘Agriculture’, for which the contaminant inputs and, hence, exposure, are variable.

On the other hand, there are good arguments for using variably acceptable ecological risks levels for different land uses, certainly with regard to Risk Management purposes (ecological soil recovery).

The soil ecosystem is an important entity for above-ground processes and above-regional (or even global) processes. The presence of plants, bushes and trees is crucial at most sites. The global significance of the soil ecosystem, related to nutrient and carbon cycling and the fixation of greenhouse gases, is important everywhere, independent of land use. Also, the cleaning function with regard to groundwater quality is crucial at most sites (and especially at agricultural and industrial sites, since generally the ‘supply’ of contaminants is highest at these sites). Without this cleaning function, the use of pesticides would even be impossible in agricultural management practices, without compromising the groundwater quality. From the perspective of the cleaning function, the inclusion of a ‘basic ecological protection level’, at every site independent of land use, is politically defensible. The possible use of such a basic protection level and the corresponding value for this protection level obviously fall into the category of a policy decision.

With regard to site-specific Ecological Risk Assessment, there are good arguments for linking ecological protection limits to land use. In other words: to link the level of acceptable damage of the soil ecosystem to the ‘ecological relevance’ of the corresponding land use. The reasoning behind this is that from a political viewpoint a small disturbance of the soil ecosystem in a nature reserve can be worse than a huge disturbance of the soil ecosystem at an industrial park (assuming that at least some basic growth of trees and plants is possible at the industrial park).

The land uses ‘Nature’ and ‘Agriculture’ have a status apart. Some continents (e.g., Antarctica) and countries (e.g., Australia, Canada, etc.) are still endowed with a pristine or almost pristine environment. For some, to allow any contamination to develop while dismantling human installations and remediating a contaminated site located in such an area is perceived as a step which, over time, will inevitably lead to the general degradation of those environments. Thus some countries have adopted policies aiming at a full remediation of those sites (e.g., Quebec Soil Protection and Contaminated Sites Rehabilitation Policy, Canada), independently of any potentially acceptable ecological risks levels. Still, application of this principle must be balanced with the appraisal of the threat to the ecosystem resulting from the techniques used to remediate the site. If the excavation of slightly contaminated soils found around rare mature trees in a park may result in the dying of the trees or, next to a salmon spawning site in a river triggers a devastating erosion process, it may be ecologically sounder to drop the remediation altogether.

Next to Nature reserves, Agriculture is the land use that is most dependent on good functioning of the soil ecosystem. Arable farming and horticulture strongly depend on a good Ecological Health, especially. Contaminated land could result in a yield reduction or even make the growth of proper agricultural crops impossible. One good reason to protect the soil quality in agricultural areas is that the soil quality impacts the amount and quality of the food that is produced. Poor quality soil produces small amounts of poor quality food. Contaminants present in the soil or pesticides applied during the growth season may be absorbed in, or adsorbed on the food. On the other hand, the intensive use of the soil during agricultural practices also threatens the soil ecosystem, for example, by using pesticides and by tillage. During the last few decades the relationship of farmers to the soil ecosystem has come under pressure. Today, a general belief in the need for more sustainable agricultural practice is gaining popularity in all developed countries, worldwide (see Section 13.8.3). Although politically debatable, the ecological relevance of land use decreases in this order: Nature, Agriculture, Residential areas (the relevance depending to a large extent on the unsealed surface area, for example gardens and other green areas), Infrastructural work and Industrial areas.

13.7.4 Secondary Poisoning and Food Web Approach

Secondary poisoning refers to a chain of feeding relationships, where each organism higher in the food chain is exposed to contaminants (mainly contaminants with a high accumulation affinity such as metals, Polycyclic Aromatic Hydrocarbons (PAHs) and other hydrophobic compounds) that have been taken up by the organism lower in the food chain. The consequence is that the organisms that are highest in such a food chain can experience *bioaccumulation*, that is, accumulation of contaminants in their tissue although these organisms often do not have a direct relationship with contaminated sites.

In fact, all above-ground animals can suffer from secondary poisoning, since herbivores consume potentially contaminated plants and carnivores consume animals that might be exposed to contaminants from soil. However, the most relevant cases of secondary poisoning relate to specific higher organisms that feed on worms (for example, badgers) and rodents (for example, birds of prey).

Although scarcely used, food web modelling (e.g., Luttik 2003) offers a tool for assessing the risks from secondary poisoning, mainly useful for site-specific Ecological Risk Assessment. Traas et al. (1996), for example, presented a secondary poisoning model for the kestrel and the barn owl. However, the resulting Ecological Risk Assessment is very sensitive to the relatively uncertain calculation of the bioaccumulation and the transfer of the contaminants within the food web (Traas 2004). Food web modelling counterbalances two of the great demerits of Ecological Risk Assessment: firstly, the focus on effects for a single species, or at best for a small number of species; secondly, the fact that differences in sensitivities to contaminants among species are often neglected. Food web approaches provide an opportunity to address the effects of contaminants at the level of an ecosystem. Alternatively, hybrid food web/population models, which combine food web models and population models, can be used for Ecological Risk Assessment at the population level (Van den Berg et al. 1998).

Food web modelling could also lead to the identification of indicator organisms that are representative for the functioning of a specific soil ecosystem. In principle, it is possible to derive Soil Quality Standards on the basis of food webs and the existing feeding relationships between the different species in the food web. This, however, requires a completely different approach in derivation of Soil Quality Standards.

The statements in the previous section on the quality of food produced in agricultural areas can be seen as an example of potential secondary poisoning due to the soil quality.

13.7.5 Wildlife Protection

In less densely populated countries where larger natural areas are present such as the US, Canada, Australia and Africa, *wildlife* is often considered as an ecological protection target. Larger animals often are vulnerable to exposure from contaminated sites, via bioaccumulation in the food chain (*secondary poisoning*, see Section 13.7.4). Therefore, mainly contaminants with a high accumulation affinity such as metals, PAHs and Polychlorinated biphenyls (PCBs) are a threat to wildlife.

Sobańska (2004), for example, measured mercury in the hair of wild boars in Poland. Almli et al. (2005) measured elevated concentrations of mercury and lead in crocodiles in the Kafue and Luangwa National Parks in Zambia. And Duffy et al. (2005) found higher mercury concentrations in the hair of free-ranging Alaskan reindeer in the Seward Peninsula, Alaska, US, than in domesticated reindeer. Wijnhoven et al. (2008) measured the concentrations in the liver and kidneys of voles, mice and mammals in a moderately contaminated flood plain in the Netherlands. Depending on the effect concentrations used, the extent of exceeding of these values was measured, mainly for shrew species.

Risk Assessment with regard to wildlife protection is generally based on measurements. Alternatively, exposure models or food-chain models can be used to calculate potential accumulation in wildlife. Fairbrother (2003) claimed that methods for assessing risk to wildlife from exposure to environmental contaminants remained highly uncertain, as empirical data required for accurate estimates of exposure or determination of toxicity thresholds were lacking. For the purpose of limiting these uncertainties, the author suggested a tiered methodology based on three multiple lines of evidence that are gathered by proceeding through a tiered approach, including:

- the concentration of contaminants in relationship to levels reported to be harmful;
- bioassays or toxicity studies to define dose-response relationships;
- field studies of population or community responses.

13.7.6 Scale and Contaminant Pattern

One important difference with Human Health Risk Assessment and Human Health Risk Management is the fact that in Ecological Risk Assessment and Risk Management the degree of ecological damage from soil contaminants is directly related to the size of the contaminated site. In addition, the Ecological Risk Assessment must often be approached for a wider area than the contaminated site alone. The role of area size in Ecological Risk Assessment has a number of reasons for it. Firstly, unlike Human Health Risk Assessment, ecological damage of an area is not equal to the sum of the damages at different contaminated sites in that area. When, for example, small areas are contaminated while the soil and the largest part of a greater area is clean, Biodiversity and Ecosystem Services are hardly threatened at all on the scale of the greater area. Alternatively, when the wider surroundings of a contaminated site are even more contaminated, the ecological value of the region is not increased that much when this (relatively small) contaminated site is remediated. Big, uncontaminated, or even virgin areas such as the arctic, however, may never be used as an excuse to trivialise the threat to the soil ecosystem at smaller contaminated sites within such an area.

Moreover, although most organisms simply reside on the spot were they came into being, many other organisms are able to migrate to the most optimal locations, mainly triggered by optimal food supply. An alternative reason for migration is to find locations (*refugia*) that are not or are less contaminated. Generally speaking, larger organisms are more mobile. The most dynamic soil organisms are earthworms (with a travelling distance of 8–10 m/year) and moles.

As a consequence, the ecological risk assessor must take into account the size of the contaminated site and the contaminant pattern when focusing on Ecological Risk Assessment.

13.7.7 Spatial Planning

In spatial planning, it is important to select the locations for a human activity on a place where the soil has the best characteristics for these activities. That results in, for example, agriculture on clean and the most fertile soil, housing in areas where soil is less fertile, but has a high bearing power, and the use of clean areas for nature and drinking water supply and not for industrial areas.

In Section 1.8.6.2, focusing on the integration of soil management with spatial planning, it was mentioned that the soil (both the upper layer and the groundwater) imposes limitations as well as opportunities for functions on the soil surface. From an ecological perspective, it is beneficial to profit from the opportunities and to deal with the limitations in spatial planning procedures. The opportunities and limitations refer to those Ecosystem Services that facilitate the growth of plants, namely, the humus formation, carbon and nutrient cycling and disease control. These Ecosystem Services are important from an agricultural and nature perspective, as well as from a decorative viewpoint related to plants, bushes and trees in gardens, parks, boulders, and industrial estates.

One example of an important ecological function to be taken into account in spatial planning is the contribution of organisms to the derivation of clean groundwater for drinking water and irrigation. This imposes preconditions for maintaining the self-cleaning capacity and the hydrological buffer functions of the soil.

In practice, ‘green functions’ should be avoided at heavily contaminated sites, especially with PAHs, Persistent Organic Pollutants and metals, since these contaminants will hamper Ecological Services for longer time periods. It must be noted, however, that adapting the land use to poor chemical soil quality, for example, sealing off highly contaminated sites in order to limit leaching – one alternative for site remediation that is maybe plausible from a practical point of view – is politically debatable.

Inversely, spatial planning should include possibilities for an optimal ecological functioning of the soil as one of the planning criteria.

13.8 Sustainability

13.8.1 Political Significance

The magic word in environmental policy today is *sustainability*. Although the term is difficult to define and many interpretations exist (see Section 1.9), sustainability in soil policy is related to a good soil quality ‘here and now’ and, in the context of this book, ‘in the wider surroundings and in the future’ as well. It is generally acknowledged that the strain on the environment, including on soil ecosystems, is so intense that its functioning for future generations cannot be taken for granted (Millennium Ecosystem Assessment 2005). And the worldwide strain on soil ecosystems is expected to increase in the future even more. Agricultural needs, for example, may demand a tripling of the amounts of nitrogen and phosphorous where agricultural practices are used in order to achieve a doubling in food production (Tilman et al. 2001). Therefore, new incentives and policies for ensuring the sustainability of Ecosystem Services will be crucial if we are to meet the demands of improving yields without compromising environmental integrity (Tilman et al. 2001). This is even more important when future generations need to be able to profit from past developments, the general idea behind the ‘cradle to cradle’ philosophy (McDonough and Braungart 2002).

Generally, sustainability is approached from an environmental, economic and social context. Environment, or ecology, is often made central to many approaches, ascribing a higher hierarchic level to it. From this perspective, sustainability primarily focuses on the environment and, hence, on the soil ecosystem (‘here and now’, and ‘in the wider surroundings and in the future’), while the provision of a good economic and social structure are important boundary conditions. Attention to sustainable development got a big push at the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro (United Nations 1992). Sustainable development, mainly for developed countries, was strongly advocated at the conference. The aim was that unsustainable patterns of production and consumption needed to be eliminated.

The association of an ecologically healthy soil and sustainability promoted the position of the soil ecosystem when defining Risk Management solutions. Remediation Objectives, for example, often include ecological criteria. Dawson et al. (2007) concluded that there has been a move towards the consideration of biological indicators for Hazard Assessment in conjunction with the remediation of contaminated soils.

13.8.2 The Benefits of Sustainability

In the early days, sustainable soil management was the logical option in agricultural practice, since humans were not able to repair land that was depleted. The only alternative that early man had was to leave the land and look for other places to grow crops or to graze their cattle. For many centuries, humans were able to live in close cooperation with natural physical soil conditions such as agricultural micro-relief, and existing rivers and ditches for drainage and irrigation. More recently, however, humans have learned to control physical soil conditions on a regional scale (starting in many countries in the nineteenth century) and learned to control the chemical soil conditions (since the twentieth century) through the use of fertilisers and pesticides. As a result, it was often believed that humans were able to allow themselves a short-term policy vis-à-vis land management. However, during the last decade or two there has been a growing awareness of the fact that short-term soil management approaches, which focus on short-term profits at the site of application only, are failing. The side effects of large scale artificial fertilizers, for example, although effective in the short-term at the sites where these have been applied, have long been underestimated. With regard to ecological effects in the long-term, accumulation of contaminants in soils hampers effective ecological functioning and application of pesticides has adverse effects on disease controlling organisms, for example. Ecological effects also may reveal themselves at other locations than those where the contaminants were applied, for example, in surface waters and groundwater due to intensive manuring.

Although many soil factors can indeed be ‘established’ through human manipulation, the general feeling is that the negative side effects are too much of a burden since they have grown significantly. Moreover, it is more and more the belief that sustainable agriculture, with the application of organic fertilizers, a minimum of pesticide use and tillage, is a preferable procedure from an economic and social perspective as well. Similar examples could be given for contaminated sites in urban areas where the ecosystems are often challenged beyond their capabilities to recover. Except for creating good conditions for above-ground vegetation, the ‘cleaning’ function of organisms is of major importance at these sites.

An extraordinary situation in which sustainable soil management proves its benefits are land transitions, that is, situations in which the land use changes. Generally, the frequency in change of land use has intensified over the last few decades in most developed countries, and probably will continue to grow. Specifically, the transition from agricultural land to nature reserves and to residential or industrial areas has often taken place in many developed countries in Europe, for example. A sound sustainable policy anticipates soil dynamics and increases change possibilities for successful land-use transition in the future. In cases of transition from agricultural land to a nature reserve, for example, agricultural soil management may have resulted in nutrient-rich soil layers that block natural developments for periods varying from decades up to a century.

13.8.3 Agriculture

Descriptions of good soil quality for agricultural purposes include many structural aspects such as ‘crumbles easily’, ‘drains well’, ‘warms up quickly’, ‘does not crust’, ‘soaks up heavy rains’, ‘stores moisture’, ‘has few clods’, ‘has no hardpan’, ‘resists erosion’, ‘allows root penetration’ (see Section 13.4.3.2 for more details). Moreover, a decent humus formation, good nutrient supply and disease control (see Section 13.4.3.3, 13.4.3.4, and 13.4.3.6, respectively) are indispensable for an optimal agricultural production. And all these characteristics may be seen as the overall result of a healthy soil ecosystem. Of course, humans can stimulate good soil structure by ploughing, either by hand or mechanically. And humans also can improve the chemical quality of the soil, through nutrient supply and the addition of lime and, indirectly, by stimulating nutrient buffering. But the soil ecosystem provides great support, working every minute of the day in each cubic centimetre of the soil. And the impressive work of the soil ecosystem in improving soil quality does not cost anything. Moreover, soil organisms will create a natural equilibrium, while human beings are forced to constantly keep up their soil management. Therefore, the soil ecosystem is to be treated with great respect in sustainable agriculture. This generally requires a focus on a fruitful co-operation between men, accepting a minimal role, and soil organisms. Most soils left alone will develop at least an acceptable soil structure 'on their own'. As an example, strategic and applied research has demonstrated that certain co-operative microbial activities can be exploited, as a low-input biotechnology, to help sustainable, environmentally-friendly, agro-technological practices (Barea et al. 2005).

13.8.4 Improving Sustainability

An important pillar of sustainability relates to the resilience and restoration of the soil ecosystem. Here, Ecological Risk Management offers possibilities. Ecosystem Services should be used as often as possible, instead of using artificial tools for soil management. Optimal conditions for Ecosystem Services, therefore, should be maintained or created, also in regard to the presence of contaminants. A good example of a sustainable process in soil is the repeated cycling of the same nutrients in soil by organisms: first, soil dwelling organisms release immobile minerals and convert them into nutrients that are available to plants. Then, when the plants die, organisms will decay the plant material and, subsequently, recycle the nutrients. In principle, this recycling will continue indefinitely and the soil organisms will be responsible for plant nutrition without any human interference.

13.9 Monitoring the Soil Ecosystem Quality

13.9.1 Indicators

Monitoring is the repetitive measurement on the same place in order to determine trends over time. Monitoring the quality of the soil ecosystem requires insight into the soil processes and the development of indicators. Ecological quality depends on chemical, physical, and biological characteristics of the soil. Chemical monitoring of soil is often performed in terms of measurements of the concentration of contaminants. In soil science, there is a good understanding about the physical characteristics of the soil, including soil type, bearing capacity, groundwater levels and so on. Not so much is known about soil biology. Many indicators exist for sub-cellular and cellular species, population and at the community level. However, application of these indicators, measurement of biological entities in a monitoring system, does not occur very often. Avidano et al. (2006) used functional (substrate utilization pattern and enzymatic activities of the soil extracts) and structural (bacterial population density and composition and structure) diversity of the soil bacterial community as indicator of Soil Health. The indicators that optimally differentiated the extent of soil remediation for a former gas works site in the UK were biomass-C, respiration, dehydrogenase activity, earthworm toxicity and mustard seed germination (Dawson et al. 2007). The authors also concluded that once robust and sensitive biological indicators were incorporated into a quantitative soil quality index, although they had different endpoints, they gave a clearer representation of Ecological Health than chemical data alone by their integration of contamination effects at a number of trophic levels. Hankard et al. (2004) investigated the performance of two earthworm biomarkers, these are, lysosomal membrane stability and the total immune activity (TIA), with the purpose to assess the impact of a range of metals and organic contaminants in soils in three industrial sites in the UK. More examples are given in Breure et al. (2005) and Rutgers et al. (2008).

Bioindicators can also be used for other purposes than assessing the impact of soil contamination. Ötvös et al. (2003), for example, used moss species as bioindicators for the impact of atmospheric deposition of several metals, in Hungary.

13.9.2 Significance of Monitoring the Soil Ecosystem Quality

When ‘sustainable Ecosystem Services’ is earmarked as a policy goal, monitoring becomes increasingly important in assessing ecological quality and in determining the effectiveness of the measures taken. Ecosystem monitoring should consist of the integrated measurement of the chemical, physical and biological characteristics of the soil. It has been applied in agriculture, for instance, in order to determine the necessity of the application of fertilizer or to determine the presence of pest organisms such as plant predatory nematodes of the fungus *Phytophtera infestans*. Monitoring is also useful for assessing the impact of soil contamination and other stress factors on the soil ecosystem (see also Section 13.5.5). In the era of sustainability, the need for ecological soil indicators has gained enormously in terms of interest.

Determination of the actual ecological quality of a soil is also necessary, for example, in site-specific Ecological Risk Assessment, where not only the extent of contamination, but also the effects of contaminants on the ecosystem are taken into account in making the decision on how to manage a site.

The Spanish regulations for assessing and managing contaminated sites, released in 2005, include direct ecological toxicity testing for the purpose of assessing the ecological effects of soil contamination (Tarazona et al. 2005). For this purpose, specific soil microcosms are in development (Fernández et al. 2005).

13.9.3 Possibilities for Monitoring

As mentioned earlier, many biological measurements are possible based on the determination of the presence or absence of organisms, or on the measurement of the activities of soil organisms. These measurements can be on the sub-cellular level (DNA, proteins), cellular level (activity of cellular processes), or on the level of species (determination of presence, absence and abundance of species), populations or communities (diversity and composition of groups of species). There has been a lot of debate on the relationship between the numbers of organisms and the quality of Ecosystem Services or Ecosystem Health. In point of fact, no universally accepted indicators exist that represent the quality of Ecosystem Services or Ecosystem Health.

Many techniques exist and are also in the process of standardisation in NEN (Dutch Standardisation Institute) and ISO (International Organization for Standardization), but application in a monitoring network is complex and expensive. In site-specific Risk Assessment, especially, these techniques are being applied, however.

One popular criterion for Ecosystem Health is a description of the community composition in comparison with the composition of a reference community. To assess the impact of contaminants on the soil ecosystem, this reference community could be a similar ecosystem under comparable conditions as at the contaminated site, but without the contamination. Although this criterion certainly has a certain relationship with Ecosystem Health, it was concluded in Section 13.5.1, however, that the unimpaired condition in which ecosystems show no influence from human activities (Ecological Integrity) is not necessarily the optimal condition for the soil ecosystem.

It is important to note that the principle of monitoring of Ecological Health is dependent on the size and complexity of the species (Beck et al. 2005). Generally, indicators related to microorganisms focus on functions such as respiration, or activity. Indicators for higher organisms, however, usually relate to the structure of the community instead of functions. One important line of research for the future will be the coupling of ecosystem structure and functioning (Breure et al. 2008).

Rutgers et al. (2008) derived ten references for good biological soil quality. The references are specific to ten combinations of land use (e.g., dairy farms, arable fields and headlands) and soil type (sand, peat, clay and loess).

13.9.4 Biological Classification Systems

Some guidebooks indirectly refer to the presence of a healthy soil ecosystem, for example, by the criterion that it ‘has a rich, earthy smell’, the unmistakable proof of the metabolisms of an almost inconceivable number of soil organisms. Some *quick* evaluations of Soil Health refer to biological aspects such as the presence of earthworms, the earthy smell, and the loose, crumbly structure of the soil (Bowman 1994).

In Europe, several indicator systems have been developed. In Denmark, for example, seven different characteristics for Soil Health are considered. Subsequently, these characteristics are coupled with Ecosystem Services-related soil ecological parameters and microbial indicators (National Environmental Research

Institute 2002). Breure et al. (2005) described two mature ecological classification systems: the Dutch Biological Indicator for Soil Quality (BISQ) and the German Soil Biological Indicator for Soil Quality (BBSK). These classification systems enable the monitoring of ecological soil quality.

13.9.4.1 Biological Indicator for Soil Quality (BISQ)

The Dutch Biological Indicator for Soil Quality (BISQ) includes information on the presence, the amounts and activities of organisms, the rate of organism-induced processes, as well as abiotic factors, as indicators for ecological soil quality (Breure et al. 2004). The actual values of these parameters are related to predefined reference values. The result can be visualized by a so-called 'amoeba' which gives the distance-to-target scores for all different factors as fan-shaped segments. In Fig. 13.4 an example of such an amoeba is given for ten Ecosystem Services for four farms in the Hoeksche Waard, the Netherlands, and compared with the national reference for arable farmland on clay (100% circle) (Rutgers et al. 2008).

In general, three different targets can be used, namely, minimal soil quality, acceptable soil quality and optimal soil quality. To determine an acceptable soil quality, a reference ecological soil quality, adapted from nature reserves or biological farms with comparable soil types and abiotic characteristics, must be defined. Another option is to make a theoretical reference, based on knowledge of the constitution of a stable or desired system. In Section 13.5.1, it was stressed that a pristine ecosystem is not necessarily an optimal ecosystem.

The BISQ is generally believed to be a comprehensive tool for monitoring soil quality, covering nearly the entire soil food web. The big drawback, however, is that it requires intensive sampling and analyses of a whole range of biological and abiotic parameters. For this reason, it has been suggested that a smaller set of parameters should be developed that need to be measured in order to achieve a better quality versus cost ratio.

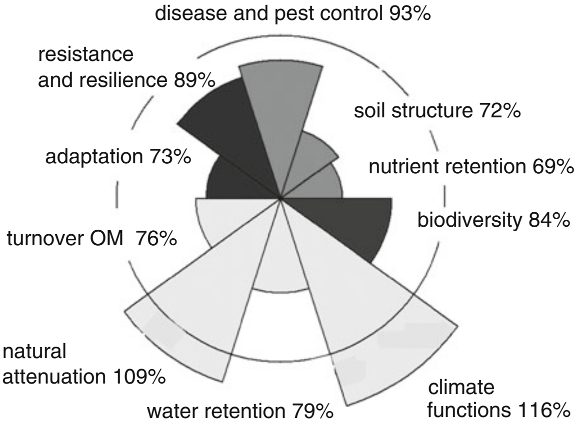


Fig. 13.4 A so-called 'amoeba', showing the distance-to-target scores for relevant Ecosystem Services as fan-shaped segments, visualising ecological soil quality (source: Rutgers et al. (2008); reproduced with permission)

References

Akhtar M, Malik A (2000) Roles of organic soil amendments and soil organisms in the biological control of plant-parasitic nematodes: a review. Bioresource Technol 74(1):35–47

Almli B, Mwase M, Sivertsen T, Musonda MM, Flåøyen A (2005) Hepatic and renal concentrations of 10 trace elements in crocodiles (*Crocodylus niloticus*) in the Kafue and Luangwa rivers in Zambia. Sci Total Environ 337(1–3):75–82

Avidano L, Gamalero E, Cossa GP, Carraro E (2006) Characterization of soil health in an Italian polluted site by using microorganisms as bioindicators, Appl Soil Ecol 30(1):21–33

The Bacteria Guide (2009) http://go-to1.com/bacteria. Retrieved 24 Feb 2009

Barea JM, Pozo MJ, Azcón R, Azcón-Aguilar C. (2005) Microbial co-operation in the rhizosphere. J Exp Bot 56:1761–1778

Barron MG, Wharton SR (2005) Survey of methodologies for developing media screening values for ecological risk assessment. Integrated Environ Assess Manage 1(4):320–332

Beck L, Römbke J, Breure AM, Mulder C (2005) Considerations for the use of soil ecological classification and assessment concepts in soil protection. Ecotoxicol Environ Safety 62(2):189–200

Bloem J, Breure AM (2003) Microbial indicators. In: Markert BA, Breure AM, Zechmeister HG (eds) Bioindicators and biomonitors. Elsevier Science, Oxford, pp 259–282

Boekhold A (2008) Ecological risk assessment in legislation on contaminated soil in The Netherlands. Sci Total Environ 406(3):518–522

Bowman G (1994) Why soil health matters. The New Farm, January 1994, http://newfarm.rodaleinstitute.org/depts/nf_classics/0905/soil_print.shtml. Retrieved 19 Mar 2009

Boivin M-E, Breure AM, Posthuma L, Rutgers M (2002) Determination of field effects of contaminants – significance of pollution-induced community tolerance. Human Ecol Risk Assess 8(5):1035–1055

Boyle M, Frankenberger WT, Stolzy LH (1989) The influence of organic matter on soil aggregation and water infiltration. J Prod Agric 2:209–299

Breure AM, Mulder Ch, Rutgers M, Schouten T, De Zwart D, Bloem J (2004) A biological indicator for soil quality. In: Proceedings from an OECD expert meeting Rome, Italy, March 2003: Agricultural impacts on soil erosion and soil biodiversity: developing indicators for policy analysis, pp 485–494

Breure AM, Mulder Ch, Römbke J, Ruf A (2005) Ecological classification and assessment concepts in soil protection, Ecotoxicol Environ Safety 62(2):211–229

Breure AM, Groot M, Eijsackers HJP (2008) System-oriented ecotoxicological research: which way to go? Sci Total Environ 406:530–537

Buckley DH, Schmidt TM (2001) The structure of microbial communities in soil and the lasting impact of cultivation. Biomed Life Sci 42(1):11–21

Callicott JB, Crowder LB, Mumford K (1999) Current normative concepts in conservation. Conservation Biol 13(1):22–35

Canadian Council of Ministers of the Environment (1999) Canadian environmental quality guidelines. Winnipeg (MB), Canada

Carlon C, Swartjes F (2007) Analysis of variability and reasons of differences. In: Carlon (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities, JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

Chapple CK (2006) Jainism and ecology: transformation of tradition. In: Gottlieb (ed) The Oxford handbook of religion and ecology. Oxford university, New York

Clauwaert P, Aelterman P, Pham TH, De Schamphelaire L, Carballa M, Rabaey K, Verstraete W (2008) Minimizing losses in bio-electrochemical systems: the road to applications, Appl Microbiol Biotechnol. doi:10.1007/s00253-008-1522-2

Conklin AR Jr (2002) Soil microorganism. Soil, Sediment Water, AEHS magazine

Commission of the European Communities (2006) Communication from the commission to the council, the European parliament, the European economic and social committee and the committee of the regions. Thematic strategy for soil protection. COM (2006)231 final, Brussels, 22.9.2006

Costanza R, d'Arge R, de Groot R, Farber S, Grasso M, Hannon B, Limburg K, Naeem S, O'Neill RV, Paruelo J, Raskin RG, Sutton P, van den Belt M (1997) The value of the world's ecosystem services and natural capital. Nature 387:253–260

Cuypers MP (2001) Bioavailability of polycyclic aromatic hydrocarbons in soils and sediments. Prediction of bioavailability and characterization of organic matter domains. Dissertation Wageningen University, Wageningen, the Netherlands

Daily GC, Söderqvist T, Aniyar S, Arrow K, Dasgupta P, Ehrlich PR, Folke C, Jansson AM, Jansson BO, Kautsky N, Levin S, Lubchenco J, Mäler K-G, Simpson D, Starrett D, Tilman D, Walker B (2000) The value of nature and the nature of value. Science 289:395–396

Dawson JJC, Godsiffe EJ, Thompson IP, Ralebitso-Senior TK, Killham KS, Paton GI (2007) Application of biological indicators to assess recovery of hydrocarbon impacted soils. Soil Biol Biochem 39:164–177

De Zwart D, Posthuma L (2005) Complex mixture toxicity for single and multiple species: proposed methodologies. Environ Toxicol Chem 24:2665–2672

De Leo GA, Levin S (1997) The multifaceted aspects of ecosystem integrity. Conservation ecology [online] 1(1):3. Available from the internet. url:http://www.consecol.org/vol1/iss1/art3/. Retrieved 19 Mar 2009

Doelman P, Jansen E, Michels M, Van Til M (1994) Effects of heavy metals on microbial diversity and activity as shown by the sensitivity-resistance index, an ecologically relevant parameter. Biol Fertil Soils 17:177–184

Duffy LK, Duffy RS, Finstad G, Gerlach C (2005) A note on mercury levels in the hair of Alaskan reindeer. Sci Total Environ 339(1–3):273–276

Dykhuizen DE (1998) Santa Rosalia revisited: why are there so many species of bacteria. Antonie van Leeuwenhoek 73:25–33

EC (2000) Directive of the European parliament and of the council of 23 October 2000 establishing a framework for Community action in the field of water policy (2000)/60/EC, Commission of the European Communities, Brussels, 23 October 2000

EC (2006) Proposal for a directive of the European parliament and of the council establishing a framework for the protection of soil and amending Directive 2004/35/EC, COM(2006) 232 final (2006)/0086 (COD), Commission of the European Communities, Brussels, 22 September 2006

Edwards CA, Bohlen PJ (1996) Biology and ecology of earthworms, 3rd edn. Chapman and Hall, London

Edwards CA, Burrows I (1988) The potential of earthworm composts as plant growth media. In: Edwards CA, Neuhauser E (eds) Earthworms in waste and environmental management. SPB Academic, the Hague, the Netherlands, pp 21–32

Efroymson RA, Will ME, Suter GW II, Wooten AC (1997) Toxicological benchmarks for screening contaminants of potential concern for effects on terrestrial plants: 1997 revision. Oak Ridge National Laboratory, Oak Ridge (TN), USA, ES/ER/TM-85/R3

Eijsackers H (2004) Leading concepts towards vital soils. In: Doelman P, Eijsackers H (eds) Vital soil, function, value and properties, developments in soil science 29, Elsevier

Eijsackers HJP, Groot M, Breure AM (2008) Upgrading system-oriented ecotoxicological research, Sci Total Environ 406(3):373–384

Emmerling Ch, Schloter M, Hartmann A, Kandeler E (2002) Functional diversity of soil organisms – a review of recent research activities in Germany. J Plant Nutr Soil Sci 165:408–420

Ernst G, Frey B (2007) The effect of feeding behavior on Hg accumulation in the ecophysiologically different earthworms *Lumbricus terrestris* and *Octolaseon cyaneum*: a microcosm experiment. Soil Biol Biochem 39(1):386–390

European Commission DG ENV (2010) Soil biodiversity: functions, threats and tools for policy makers, contract 07.0307/2008/517444/ETU/B1, final report, February 2010, Bio Intelligence Service S.A.S. in association with IRD and NIOO

European Environment Agency (2008) EEA briefing 2008/02, European environment agency, TH-AM-08–002-EN-N, Copenhagen

Fairbrother A (2003) Lines of evidence in wildlife risk assessments. Human Ecol Risk Assess 9(6):1475–1491

Fernández MD, Cagigal E, Milagrosa Vega M, Urzelai A, Babín M, Pro J, Tarazona JV (2005) Ecological Risk assessment of contaminated soils through direct toxicity assessment, Ecotoxicol Environ Safety 2(2):174–184

Fishwick S (2004) Soil screening values for use in UK. Ecological risk assessment. Environment Agency

Gottlieb RS (2006) Religion and ecology – what is the connection and why does it matter? In: Gottlieb RS (ed) The Oxford handbook of religion and ecology. Oxford university, New York

Griffiths BS, Ritz K, Bardgett RD, Cook R, Christensen S, Ekelund F, Sørensen S, Bååth E, Bloem J, de Ruiter P, Dolfing J, Nicolardot B (2000) Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions: an examination of the biodiversity-ecosystem function relationship. Oikos 90:279–294

Griffiths BS, Ritz K, Wheatley R, Kuan HL, Boag B, Christensen S, Ekelund F, Sørensen SJ, Muller S, Bloem J (2001) An examination of the biodiversity–ecosystem function relationship in arable soil microbial communities. Soil Biol Biochem 33(12–13):1713–1722

Haas D, Défago G (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads, Nature Reviews Microbiology, AOP, published online 10 March 2005; doi:10.1038/nrmicro1129

Hankard PK, Svendsen C, Wright J, Wienberg C, Fishwick SK, Spurgeon DJ, Weeks JM (2004) Biological assessment of contaminated land using earthworm biomarkers in support of chemical analysis, Sci Total Environ 330(1–3):9–20

Hunt HW, Coleman DC, Ingham ER, Ingham RE, Elliott ET, Moore JC, Rose SL, Reid CPP, Morley CR (1987) The detrital food web in a shortgrass prairie. Biol Fertility Soils 3(1–2):57–68

Ingham ER (2000a) The soil food web, chapter 1. In: Tugel AJ, Lewandowski AM, Happe-vonArb D (eds) Soil biology primer. rev. ed. Ankeny, Soil and Water Conservation Society, Iowa

Ingham ER (2000b) Soil fungi, chapter 4. In: Tugel AJ, Lewandowski AM, Happe-vonArb D (eds) Soil biology primer. rev. ed. Ankeny, Soil and Water Conservation Society, Iowa

Jager T, den Hollander HA, van der Poel P, Rikken MGJ, Vermeire T (2001) Probabilistic environmental risk assessment for dibutylphthalate (DBP). Hum Ecol Risk Assess Int J 7(6):1681–1697

Jensen J, Mesman M (2006) Ecological risk assessment of contaminated land, Decisions support for site specific investigations. Report number 711701047, National Institute for Public Health and the Environment, Bilthoven, the Netherlands

Kapustka LA, Landis WG (1998) Ecology: the science versus the myth. Human Ecol Risk Assess 4(4):829–838

Koster M, De Groot A, Vijver M, Peijnenburg W (2006) Copper in the terrestrial environment: verification of a laboratory-derived terrestrial biotic ligand model to predict earthworm mortality with toxicity observed in field soils. Soil Biol Biochem 38:1788–1796

Lackey RT (2001) Values, policy, and ecosystem health. Bioscience 51(6):437–443

Lancaster J (2000) The ridiculous notion of assessing ecological health and identifying the useful concepts underneath. Human Ecol Risk Assess: Int J 6(2):213–222

Langmaack M, Wiermann C, Schrader S (1999) Interrelation between soil physical properties and Enchytraeidae abundances following a single soil compaction in arable land. J Plant Nutr Soil Sci 162(5):517–525

Lindahl BD, Ihrmark J, Boberg J, Trumbore SE, Högberg P, Stenlid J, Finlay RD (2007) Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. New Phytol 173:611–620

Lomako J, Lomako WM, Whelan WJ (2004) Glycogenin: the primer for mammalian and yeast glycogen synthesis. Biochim Biophys Acta 1673:45–55

Luttik R (2003) Risk assessment scheme for the impact of plant protection products on birds and mammals. Dissertation University of Leiden, the Netherlands

Madsen EL, Winding A, Malachowsky K, Thomas CT, Ghiorse WC (1992) Contrasts between subsurface microbial communities and their metabolic adaptation to polycyclic aromatic hydrocarbons at a forested and an urban coal-tar disposal site. Microb Ecol 24:199–213

Martino E, Cerminara S, Prandi L, Fubini B, Perotto S (2004) Physical and biochemical interactions of soil fungi with asbestos fibers. Environ Toxicol Chem 23(4):938–944

McDonough W, Braungart M (2002) Remaking the way we make things. Cradle to cradle. North Point Press USA Edition ISBN10:0-86547-587-3 | ISBN13:9780865475878

Mihail JD, Bruhn JN (2005) Foraging behaviour of *Armillaria rhizomorph* systems. Mycol Res 109:1195–1207

Millennium Ecosystem Assessment (2005) Ecosystems and human well-being: biodiversity synthesis. World Resources Institute, Washington, DC

Ministry of VROM (2008) Soil remediation circular 2006, as amended on 1 October 2008

Mueller GM, Schmit JP (2006) Fungal biodiversity: what do we know? What can we predict? Biodivers Conserv 16:1–5

Mulder Ch, Cohen JE, Setälä H, Bloem J, Breure AM (2005) Bacterial traits, animals' body mass and numerical abundance in the detrital soil food web of Dutch agricultural grasslands. Ecol Lett 8:80–90

Mulder Ch (2006) Driving forces from soil invertebrates to ecosystem functioning: the allometric perspective. Naturwissenschaften 93:467–479

National Environmental Protection Council (2003) Guideline on the investigation levels for soil and groundwater, Canberra, Australia. National Environmental Protection Measure 1999, Schedule B(1)

NYSDEC (1998). Ambient water quality standards and guidance values and groundwater effluent limitations. New York State Department of Environmental Conservation, Albany, NY, USA

Oregon Department of Environmental Quality (2001). Guidance for ecological risk assessment, level ii screening level values, Portland (OR), USA

Ötvös E, Pázmándi T, Tuba Z (2003) First national survey of atmospheric heavy metal deposition in Hungary by the analysis of mosses. Sci Total Environ 309:151–160

Perotto S, Bonfante P (1997) Bacterial associations with mycorrhizal fungi: close and distant friends in the rhizosphere, Trends Microbiol 5:496–501

Peijnenburg WJGM, Posthuma L, Eijsackers HJP, Allen HE (1997) A conceptual framework for implementation of bioavailability of metals for environmental management purposes. Ecotoxicol Environ Safety 37:163–172

Posthuma L, Suter GW, Traas TP (2002) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL

Posthuma L, Hogervorst RF, Van Straalen NM (1992) Adaptation to soil pollution by cadmium excretion in natural populations of *Orchesella cincta* (L.) (Collembola). Archiv Environ Contamin Toxicol 22(1):146–156

Rabaey K, Rodríguez J, Blackall LL, Keller J, Gross P, Batstone D, Verstraete W, Nealson KH (2007) Microbial ecology meets electrochemistry: electricity-driven and driving communities. ISME J 1:9–18

Rapport DJ (1995) Ecosystem Health: exploring the territory. Ecosystem Health 1(1):5–13

Religion and Ecology (2009) www.religionandecology.org. Retrieved 18 Feb 2009

Römbke J, Breure AM, Mulder Ch, Rutgers M (2005) Legislation and ecological quality assessment of soil:implementation of ecological indication systems in Europe. Ecotoxicol Environ Saf 62(2):201–210

Rutgers M (2008) Field effects of pollutants at the community level – experimental challenges and significance of community shifts for ecosystem functioning. Sci Total Environ 406(3): 469–478

Rutgers M, Mulder C, Schouten AJ, Bloem J, Bogte JJ, Breure AM, Brussaard L, De Goede RGM, Faber JH, Jagers op Akkerhuis GAJM, Keidel H, Korthals GW, Smeding FW, Ter C, Van Eekeren N (2008) Soil ecosystem profiling in the Netherlands with ten references for biological soil quality, RIVM report 607604009, RIVM, Bilthoven, the Netherlands

Sláviková E, Vadkertiová R (2003) Culture collection of yeasts. J Basic Microbiol 43(5):430–436

Sobańska MA (2004) Wild boar hair (Sus scrofa) as a non-invasive indicator of mercury pollution. Sci Total Environ 339(1–3):81–88

Sokhn J, De Leij FAAM, Hart TD, Lynch JM (2001) Effect of copper on the degradation of phenanthrene by soil micro-organisms. Lett Appl Microbiol 33(2):164–168. doi:10.1046/j.1472–765x.2001.00972.x

Spurgeon DJ, Sandifer RD, Hopkin SP (1996) The use of macroinvertebrates for population and community monitoring of metal contamination: indicator taxa, effect parameters and the need for a soil invertebrate prediction and classification scheme (SIVPACS), Bioindicator Syst Soil Pollut 10:96–110

Stams AJM, de Bok FAM, Plugge CM, van Eekert MHA, Dolfing J, Schraa G (2006) Exocellular electron transfer in anaerobic microbial communities, Environ Microbiol 8:371–382

Stevenson FJ, Cole MA (1999) Cycles of soil: carbon, nitrogen, phosphorus, sulfur, micronutrients, 2nd edn. Wiley, New York, NY

Struijs J, van de Meent D, Peijnenburg WJGM, van den Hoop MAGT, Crommentuijn T (1997) Added risk approach to derive maximum permissible concentrations for heavy metals: how to take natural background levels into account. Ecotoxicol Environ Safety 37:112–118

Sullivan P (2004) Sustainable soil management, ATTRA (national sustainable agriculture information service; formerly known as appropriate technology transfer for rural areas), National sustainable agriculture information service, ATTRA Publication #IP027/133, ATTRA, Fayetteville, USA

Swartjes FA (1999) Risk-based assessment of soil and groundwater quality in the Netherlands: standards and remediation urgency. Risk Anal 19(6):1235–1249

Swartjes FA, Carlon C, de Wit NHSM (2008) The possibilities for the EU-wide use of similar ecological risk-based soil contamination assessment tools, Sci Total Environ 406(3): 523–529

Takeda N, Kalam Azad MA, Ishikawa K, Yoshimura T, Fujii J, Kawasoe S (2007) Development of soil activating agents for agriculture using bakuhan-seki, J Food, Agric Environ 5(3–4): 363–367

Tarazona JV, Fernandez MD, Vega MM (2005) Regulation of contaminated soils in Spain: a new legal instrument. J Soils Sediments 5(22):121–124

Texas Natural Resource Conservation Commission (2001) Guidance for conducting ecological risk assessments at remediation sites in Texas, Austin (TX), USA

Tilman D, Wedin D, Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature 379:718–720

Tilman D, Fargione B, Wolff C, D'Antonio A, Dobson R, Howarth D, Schindler WH, Schlesinger D, Simberloff D, Swackhamer D (2001) Forecasting agriculturally driven global environmental change. Science 292:281–284

Torsvik V, Goksøyr J, Daae FL (1990) High diversity of DNA of soil bacteria. Appl Environ Microbiol 56:782–787

Traas TP, Luttik R, Jongbloed RH (1996) A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators I. Model description and uncertainty analysis. Ecotoxicol Environ Safety 34(3):264–278

Traas TP (2004) Food web models in ecotoxicological risk assessment. Dissertation IRAS, Utrecht University, Utrecht, the Netherlands

UNEP Convention of Biological Diversity (2010) http://www.cbd.int/2010/countries. Retrieved 31 Mar 2010

United Nations Conference on Environment and Development (UNCED) (1992) Report of the United Nations conference on environment and development. Rio de Janeiro, 3–14 June 1992.

Annex I: Rio declaration on environment and development. A/CONF.151/26 (vol I). Distr. Genral, 12 Aug 1992

US Environmental Protection Agency (1999) Ecological screening levels for RCRA Appendix IX hazardous constituents. Region V, Washington DC

US Environmental Protection Agency (2001) Supplemental guidance to RAGS: region IV bulletins, ecological risk assessment, Originally published November 1995

US Environmental Protection Agency (2003) Ecological soil screening level guidance, Washington DC

UNEP (2005) Millennium Ecosystem Assessment, www.millenniumassessment.org

Van den Berg MMHE, Tamis WLM, Van Straalen NM (1998) Current advantages from Europe: the food web approach in ecotoxicological risk assessment. Human Ecol Risk Assess 4(1):49–55

Van Paassen LA (2009) Biogrout, ground improvement by microbial induced carbonate precipitation. Dissertation Technical University Delft, the Netherlands

Van Straalen NM, Denneman CAJ (1989) Ecotoxicological evaluation of soil quality criteria. Ecotoxicol Environ Safety 18:241–251

Van Straalen NM, van Gestel CAM (2008) A stress ecology framework for comprehensive risk assessment of diffuse pollution. Sci Total Environ 406(3):479–483

Van der Wal A (2007) Soils in transition: dynamics and functioning of fungi. PhD Thesis Leiden University, The Netherlands

Verbruggen EMJ, Posthumus R, Van Wezel AP (2001) Ecotoxicological serious risk concentrations for soil, sediment and (ground)water: updated proposals for first series of compounds, RIVM report 711701020, RIVM, Bilthoven, the Netherlands

Vijver MG (2005) The ins and outs of bioaccumulation. PhD Thesis Free University Amsterdam, The Netherlands

Wallace KJ (2007) Classification of ecosystem services: problems and solutions. Biol Conservation 139(3–4):235–246

Wan JHC, Wong MH (2004) Effects of earthworm activity and P-solubilizing bacteria on P availability in soil. J Plant Nutr Soil Sci 167(2):209–213

Weeks JM, Comber SDW (2005) Ecological risk assessment of contaminated soil. Mineral Mag 69(5):601–613

Westra L (1998) The ethics of integrity. In: The land ethic: meeting human needs for the land and its resources. Society of American Foresters, Bethesda, MD, pp 31–44

Whitman W, Coleman D, Wiebe W (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci USA 95(12):6578–6583

Wijnhoven S, Leuven RSEW, Van der Velde G, Eijsackers HJP (2008) Toxicological risks for small mammals in a diffusely and moderately polluted floodplain. Sci Total Environ 406(3):401–406

Wright JF (2000) Assessing the biological quality of fresh waters: RIVPACS and other techniques, pp 1–24

Youn-Joo A (2004) Soil ecotoxicity assessment using cadmium sensitive plants. Environ Pollut 127(1):21–26

Zorn MI, van Gestel CAM, Eijsackers HJP (2008) Metal redistribution by surface casting of four earthworm species in sandy and loamy clay soils. Sci Total Environ 406(3):396–400

Chapter 14
Ecological Risk Assessment of Diffuse and Local Soil Contamination Using Species Sensitivity Distributions

Leo Posthuma and Glenn W. Suter

Abstract Ecological Risk Assessment related to soil contamination requires a conceptual framework and practical tools to support Risk Management. The conceptual framework is provided by the Risk Assessment paradigm, which means that risks are assessed based on an Exposure Assessment and an Effect Assessment step. Current practical tools to appraise soil quality by Ecological Risk Assessment are: (1) comparison of soil contaminant concentrations to ecological Soil Quality Standards; (2) quantification of ecological risks by modeling; (3) quantification of impacts in bioassays or in field monitoring; and (4) quantification of ecological risks by weight-of-evidence approaches. The present chapter concerns the theory and practices of Effect Assessment and risk modeling using *Species Sensitivity Distributions* (SSDs), and similar *Functional Sensitivity Distributions (*FSDs). SSD- and FSD-based Risk Assessment outputs are used for the appraisal of soil quality, soil protection and the management of (slightly and highly) contaminated sites, for both the upper soil and the groundwater. For the appraisal of soil and soil protection, one can derive *Hazardous Concentrations* (HCs) for individual contaminants, which are estimates of the concentration of a chemical that would affect a defined fraction of species. Likewise, one can derive *Hazard Potentials* (HPs) for contaminated soil samples, which represent effects levels for a certain fraction of the tested soil species when exposed in such a soil. This chapter introduces the theory of SSDs and illustrates the types of practical applications of SSD-based effect and risk models in all four of the aforementioned types of tools. Since Risk Assessment requires assessments of exposure as well as effects, the chapter also discusses Exposure Assessments for SSDs. Practical software models and database tools are described, to support easy application of SSDs in practice. Through the examples, the reader is informed on a multitude of useful options for SSD-based assessment. SSD-modeling is versatile, and can be of use to a range of soil contamination problems, from diffuse contamination in large areas to local contamination hot spots.

L. Posthuma (✉)
National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: leo.posthuma@rivm.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_14,

Contents

14.1 Aims of this Chapter and Readers Guide

This chapter aims to present and illustrate the theory and practices of using *Species Sensitivity Distribution* (SSD) modeling in the ecological effect and Risk Assessment of soil contamination.

SSDs have been used since the late 1980s in various legislations to derive environmental quality standards for water, sediment and soil. More recently, the approach has been used for other purposes including site-specific Risk Assessments of contaminated sites and various kinds of risk ranking. Risk ranking is needed when quality standards are exceeded or might be exceeded. In such cases, exceedance may occur at many sites (e.g., candidate remediation sites), they may involve different areas, or different risks are suspected (e.g., from pesticide uses). In such cases, ranking of risks amongst cases in large case loads, or for alternative management scenario's, may be of help in focusing on (cost) effective Risk Management strategies.

SSDs are Effect Assessment models, and are never applied alone. The Risk Assessment paradigm (Fig. 14.1) requires that appropriate Exposure Assessment models be used with the SSD or other effect model. This chapter thus contains appropriate Exposure Assessment tools when needed.

This chapter first presents the conceptual background for interpreting SSDs along with a description of the motives for protecting living soils and a realistic example of observed mixture effects in a soil pollution gradient. They set the scene for the remainder of the contents and examples of the chapter. We think that the reader should be acutely aware of the variance of realistic responses of soil ecosystems in the field in order to understand the strengths and limitations of SSD-modeling.

Technical and practical issues in using SSDs are subsequently presented. Some software and database references are described, to enable easy use. The practical usefulness of SSDs is illustrated and proven with examples from current practices. The validation studies and the practical examples may be of most interest to soil appraisal practitioners, since the section on validation shows how SSD-output is ground-truthed, and the section on practical uses shows the versatility and benefits of SSD-based risk modeling. The examples of use mostly concern the Netherlands.

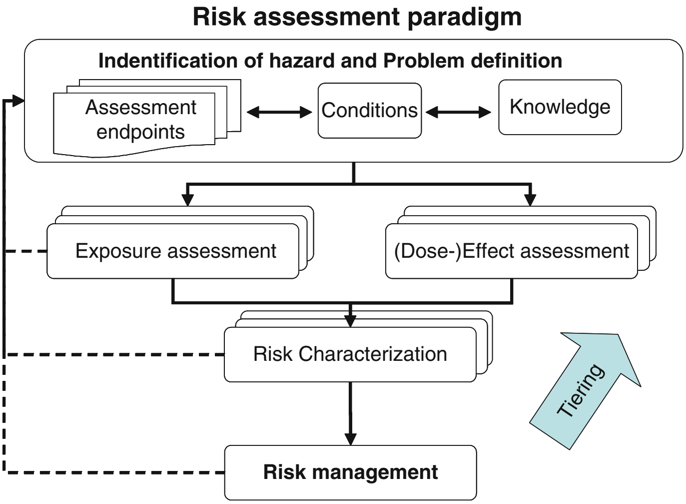


Fig. 14.1 The Risk Assessment paradigm: Exposure Assessment and Effect Assessment are used together to get insights in risks. Different methods can be applied, and those are usually organized in a tiered scheme. SSDs are the key Effect Assessment method addressed in this chapter

This is because most of the research groups that have been exploring uses of SSDs beyond the "older" use of standard setting are Dutch. Though the set of examples is thus biased, we think we can help readers to imagine how SSDs can be used to solve their own Risk Management problems. The "newer" uses of SSDs appears versatile and a useful help in Risk Management of (slightly) contaminated soils.

14.2 Soil Protection Motives and Impacts of Non-Protection

14.2.1 Protecting Living Soil – Motives

Soils are vital entities for human welfare, as described in Swartjes et al. (Chapter 10 of this book). And soils, to the contrary, are not dead. Vital soils perform a range of functions (called Ecosystem Services), ranging from global biogeochemical cycles (e.g., as carbon sinks) to prevention of flooding by regional water retention and food production. Almost all these functions require a huge and diverse community of soil organisms. Vital soils – in short – are vital for both man and nature, and enable the sustainable use of their services by mankind.

Thus, vital soils – full of living entities like bacteria, fungi, nematodes, mites, springtails, and earthworms – are generally considered worthy of protection. And, as a specific aspect of environmental protection, soil is a functionally "slow" and practically immobile compartment, as compared to the aquatic and air compartments. Impacts of stressors like contaminants often appear slowly, and often disappear equally slowly. Moreover, soil contamination is relatively immobile and confined to specific (sometimes large) areas as compared to water contamination. The “slowness” of soil can be considered an extra motive for applying the Precautionary Principle (UNESCO 2005) in soil protection.

14.2.2 Protecting Living Soil – Handling Diverse Stressor Responses

When – as a consequence of protection goals – policy makers decide that soil protection should be of local, regional, national or international importance, soil protection frameworks and associated tools are needed. The same holds for Risk Management and remediation of soils that have been contaminated by past human activities or by natural causes, like metal ore deposits that reach into the upper soil.

Given a policy decision to protect and manage soils, the issue of soil diversity arises. Soils can be contaminated with vast numbers of contaminants and mixtures. They differ vastly in physico-chemical properties (dry to wet, acid to alkaline, sandy to clayey, low to high organic content, et cetera). They can have vastly different contamination histories. They contain a vast variety of biotic species assemblages, the compositions of which reflect the cumulative influences of climate, soil composition, biogeography and evolution. As a result, we ask:

> How can we appraise, protect and manage vital soil systems when facing the threats from contaminants in this context of vast diversity and variability?

This chapter suggests an operational way to answer that question. That way is:

> Model the ecological risks of contaminants, or contaminants mixtures, over a range of concentrations from very low to very high.

Although risk modeling includes an array of empirical statistical modeling approaches and mathematical simulations, such as models of organism bioenergetics, population demographics, food web transfers, and ecosystem energetics, this chapter is mostly restricted to *sensitivity distributions*, distributions of observed sensitivities of species or of ecosystem functions.

Sensitivity distributions have been adopted (or are being discussed) in various environmental regulations worldwide, including chemicals regulation, water quality assessment, and the assessment of plant protection products. See, as examples, the technical guidance documents related to Stephan et al. (1985), OECD (1992), U.S. EPA (1995), EU (1997), EU (2000), and EU (2003). A more detailed overview of the implementation of sensitivity distributions in environmental regulations is presented in Posthuma et al. (2002b).

14.2.3 Field Effects of Soil Contamination in a Pollution Gradient

The use of the basic concept of sensitivity distributions in Risk Assessment is not at variance with observed field effects of contaminant exposure; field effects along pollution gradients clearly show that species indeed differ in their responses to pollution. One example is shown here, to illustrate the real suite of responses that may occur in an exposure situation. We consider this example key for understanding the remainder of the chapter.

Figure 14.2 shows which simultaneous responses to contaminant exposure may occur in soils. Responses are expressed as numbers of biological taxa and as activity parameters, respectively, along a copper (+ cadmium + zinc + lead) contamination soil gradient near a metal smelter in Sweden, a situation with limited influences of other stressors (Tyler 1984). Note that the graph depicts response functions in relation to total metal concentrations, while it is likely that only part of the total concentration is available for uptake due to ageing and sorption processes. Due to the relative homogeneity of the gradient soils with respect to metal binding properties, we expect that the response graphs would show similar patterns when they are expressed on the basis of metal concentrations available for uptake in the different species.

As illustrated in the figure, risk modeling using distributions of sensitivities may focus on *two* types of impacts, i.e., *structural* and *functional* impacts (the latter relating to the concept of Ecosystem Services). When present in high concentrations, one or more contaminants might impact one or more of the species that make up the vital soil. These contaminants may thereby affect the *structural integrity* of the species assemblage in the soil system. Some species are decreasing in abundance (until they are lost), while others (opportunistic species) flourish, until they are also affected at higher concentrations. The increase of such species at moderate exposures may be

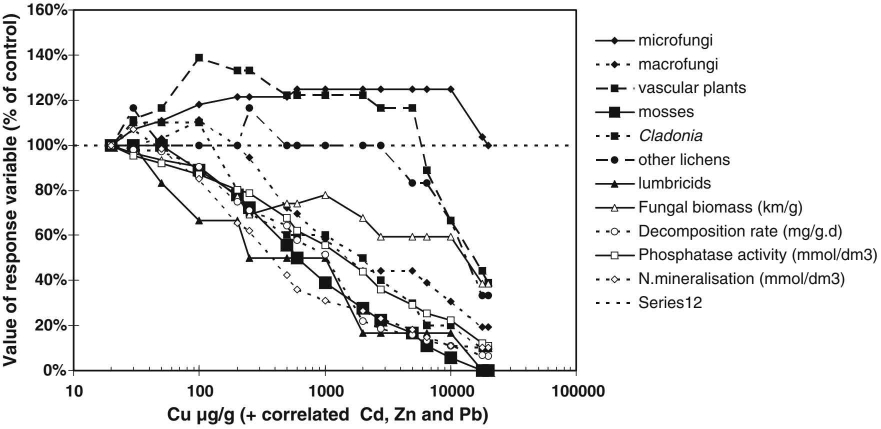


Fig. 14.2 Biotic responses as functions of copper concentration along a multiple metal gradient in the field, recalculated to 100% performance of each of the variables at the clean end of the gradient (*broken horizontal line*). This figure illustrates the application of the sensitivity distribution approach to various ecological impacts in contaminated soils. Both *structural* (*black*) and *functional* (*white*) aspects of the vital soil system change when contaminant concentrations increase. Note the opportunistic increase of some parameters (e.g., macro- and microfungi) at intermediate concentrations. The graph is redrawn from the data of Tyler (1984) for a metal smelter in a homogeneous forested area

a consequence of (relatively) low sensitivity in combination with changed ecological interactions. For example, a prey species may increase in abundance when its more sensitive predator declines. Simultaneously, some of the key *functional characteristics* of the soil system may be affected. When species decrease or increase in abundance, their contributions to the cycling of nutrients and energy also change. This eventually shows up as functional changes.

It has been demonstrated frequently that structural changes precede functional changes, especially when functional changes are considered at the level of changes in the overarching Ecosystem Services of the soil (like decomposition efficiency), and not on a basic functional response like a change in activity of a single enzyme. In other words: it can be assumed, based on available evidence, that sufficient protection of structure also would imply sufficient protection of Ecosystem Services, so that Risk Assessments could focus on the former. If one doubts this assumption, one can consider both types of effects, and appraise the soil by using the most appropriate or most sensitive endpoint. In the Netherlands, protective soil quality standards are derived in this way: preliminary standards are derived for both structure and function, and the most sensitive endpoint is used to set the regulatory standard (that then aims to protect both).

14.2.4 From Field Effects to SSD Modeling

Each species and each functional trait seems to respond in a typical way to contaminant exposure, and the responses also depend on the contaminant, the site characteristics and the other parts of the community. Given the vast variety of impact types and magnitudes, it seems hardly possible to define scientifically sound ways to derive an appraisal method that can be used in daily practice for protecting and managing vital soils threatened by contaminants. Nonetheless, instead of focusing on the apparent difficulties, one can focus on the very fact that "*All animals are unequal*" in their sensitivity to contaminants (Fig. 14.2), and that this variation per se might be a good basis for Risk Assessment. As eloquently stated:

> Variability is 'noise' for physicists, but variation is the key issue of the research of ecologists, for them, variation is 'music' (Joosse-van Damme 1984).

Instead of considering this variety as a nuisance, it is the dissimilarity in response patterns that is a valuable key fact of life: *without variability, all species would be killed at a specific concentration of a contaminant in their environment*! And it is this "music" that was recognized in the mid-eighties of the last century as a key option for appraisal and protection of biotic communities. As a result, sensitivity distribution modeling was triggered.

On both sides of the Atlantic Ocean, in the United States of America and Europe, scientists developed the modeling concept of SSDs. Key inventors were Klapow and Lewis (1979) for California, and especially Hansen, Mount and others for the federal U.S. government (U.S.EPA 1978a, b), and Kooijman (1987) and Van Straalen and

Denneman (1989) for Europe (especially the Netherlands). Glenn Suter coined the term *Species Sensitivity Distributions* for this type of modeling in OECD (1992).

The SSD method was based on the insight:

> We can see that the species sensitivity (LC50 or LD50) [that is: measures of sensitivity] distributes itself in a rather consistent way for most chemicals. The distribution resembles a lognormal one. Thus, each species we test is not representative of any other species, but is one estimate of the general species sensitivity (Mount 1982, as cited in Suter (2002)).

The SSD-modeling concept is based on the fact that species appear to differ in sensitivity to the same contaminant (like the species in Fig. 14.2). SSDs *do not explain* these differences, and *do not attempt to simulate soil ecology*. They just represent sensitivity differences to toxicants in a useful way for environmental protection and appraisal (see below). The same holds for the functional traits, in regard to Ecosystem Services, resulting from the activities of all the species. For these distributions the term Functional Sensitivity Distributions has been coined (FSDs) by Posthuma et al. (2002b). In the remainder of this chapter, we use “SSD” to refer to both SSDs and FSDs.

14.3 SSD Modeling and Practical Needs

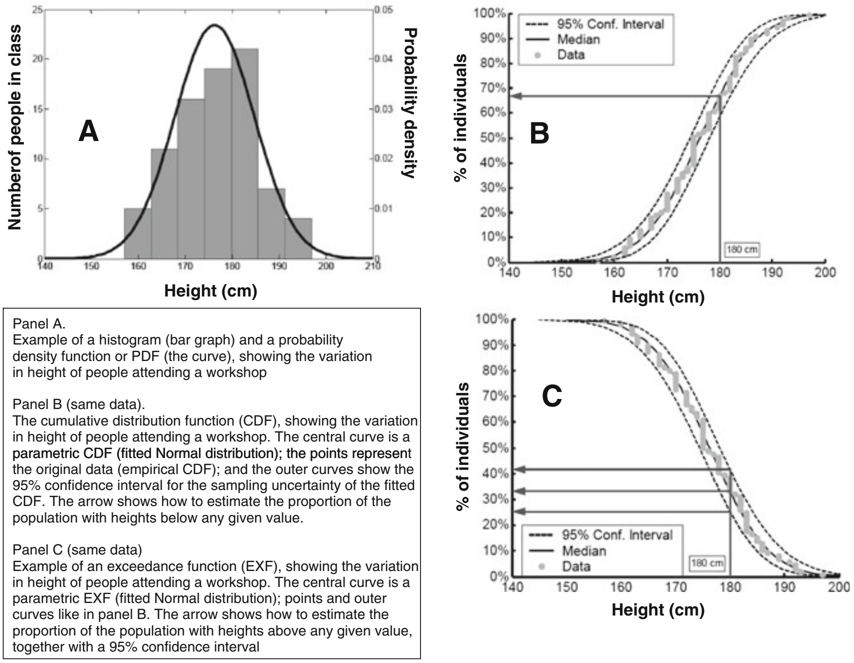
14.3.1 Basics of Distribution Modeling as an Assessment Approach

SSDs are distributions of data, in this case of laboratory data from ecotoxicity studies. These are studies in which chosen species are tested in controlled exposures to a suite of concentrations of the chosen compound. Such studies yield estimates of the sensitivity of different species to that compound, and the compilation of such data provides the input for SSD modeling. *SSD models are thus descriptive, statistical models of species sensitivity data which were collected in controlled exposures*.

Distributions of data, such as sensitivity data, are at the heart of statistics. Anyone with a basic scientific training knows concepts like the mean and the variance of a data set. SSDs and FSDs thus *describe* the variability of sensitivities across species or functions *using extremely basic statistical concepts*.

A very useful and illustrative introduction to the statistics and interpretation of species sensitivities in relation to contaminant management is provided by EUFRAM (2006). This report provides a basic explanation and visualization of distribution statistics in terms of the distribution of the variation in height of people attending a meeting (see Fig. 14.3), as follows (adapted for this chapter):

- panel A shows the raw height data summarized as a bar diagram and as a Normal probability density distribution derived from these data;
- panel B and Panel C show the same data, now as Cumulative Distribution Function and an Exceedance Function, illustrating how to estimate the fraction of the visitors’ population with heights below (B) or above (C) for any given value.



Panel A.
Example of a histogram (bar graph) and a probability density function or PDF (the curve), showing the variation in height of people attending a workshop

Panel B (same data).
The cumulative distribution function (CDF), showing the variation in height of people attending a workshop. The central curve is a parametric CDF (fitted Normal distribution); the points represent the original data (empirical CDF); and the outer curves show the 95% confidence interval for the sampling uncertainty of the fitted CDF. The arrow shows how to estimate the proportion of the population with heights below any given value.

Panel C (same data)
Example of an exceedance function (EXF), showing the variation in height of people attending a workshop. The central curve is a parametric EXF (fitted Normal distribution); points and outer curves like in panel B. The arrow shows how to estimate the proportion of the population with heights above any given value, together with a 95% confidence interval

Fig. 14.3 By creating a distribution (either a bell-shaped probability density model, **panel a**) or either of the two possible sigmoid cumulative models (**panel b** and **c**), one can summarize the key characteristics of a data set, and derive consequences of certain conditions imposed on the entities composing this data set. From EUFRAM (2006)

The fitted models allow for summary of the raw data in a few parameters, like the mean and the variance for the Normal distribution.

If this particular "sample of visitors" were gathered in a room with a door height of 180 cm, the most basic interpretation (based on raw data on height of people and the door) is that exactly those people with lower height could safely leave the room without hitting the top of the door frame. From the summary models, one can calculate that 68% of the people would be safe from hitting the top of the door frame when leaving the room (panel B), or that 32% would be unsafe (panel C). The summary models can be used in a subsequent workshop to predict the people "at risk". In such a case, the representativity of the measured sample of people is relevant, as well as the sample size. When the sample size increases, the uncertainty of the prediction is commonly reduced. This reduction of statistical uncertainty is captured in the confidence intervals shown in panels B and C. What remains is the uncertainty that the sampled people are not representative for assessing the risks for participants of other workshops in the same room.

The sigmoidal cumulative distribution, Panel B, was originally adopted to represent the SSD concept in ecotoxicology mostly in Europe (other authors prefer a linear probit model). Using a curve like Panel B, a community could be assumed protected against adverse effects of contaminant exposure by deriving an estimated maximum allowable (or trigger) soil concentration (instead of the door height) X that would protect a certain chosen (high) fraction of species (Y). The maximum allowable concentrations would differ between compounds, given a chosen Y-value, due to differences in ecotoxicity of the particular compound, i.e., the position of the model distributions of panels A, B and C shifts on the X-axis, depending on a given compound's ecotoxicity. This can be done on any chosen risk limit, a maximum allowable risk level on the Y-axis. Another use is also possible: to estimate the fraction of species (un)protected, given a certain environmental concentration. These "inverse" and "forward" uses were both proposed as formats of possible SSD use in the early stage of SSD-theory development by one of the European inventors of the approach, Professor Nico van Straalen (Van Straalen and Denneman 1989). In that paper, the fifth percentile of the distribution of sensitivities amongst soil species, based on No-Effect type input data, was chosen (Y-axis) as the maximum tolerable risk (MTR) level (whereby the choice for NOECs as input data and the fifth percentile of the chosen model were all policy choices), yielding the so-called *HC5* of a compound, the Hazardous Concentration for 5% of the species (X-axis), more precisely: 5% of the tested species. A low HC5 implies a high ecotoxicity of a compound.

Since the invention of SSDs, focus has been historically mostly on the "inverse" use of SSDs, to set environmental quality standards like standards based on HC5. Within that use, focus was most on the HC5. The "forward" use – not limited to a chosen maximum tolerable risk level – gained attention only after finding out that the use of protective standards alone was insufficient to solve existing (mixture) contamination problems.

14.3.2 Two Practical Needs and Two Useful SSD Applications

Suter and Cormier (2008) distinguished two main types of practices in Ecological Risk Assessment:

1. Criterion Risk Assessment;
2. Conventional Risk Assessment.

The "forward" and "inverse" use of the SSD model fit well with the different types of Risk Assessment needed in practice.

Criterion (or according to the terminology in this book: Soil quality standard) Risk Assessment begins with a contaminant and the policy-based selection of a type and level of effect that is not to be exceeded, and then uses the exposure-response model to eventually estimate a level of exposure that constitutes the criterion (for example, the HC5 and the formal soil quality standard). In this assessment, the choice of an effect metric (Y-value) that is consistent with the selected environmental policy goal (the protection endpoint, and its level of protection) is key. In soil protection and appraisal, the policy goal usually relates to undisturbed species diversity or undiminished Ecosystem Services, that is: endpoints at the level of the whole species assemblage. This fits well with the HC-concept.

Conventional Risk Assessments begin with a current or expected source of a contaminant or a contaminated site, which is used to estimate exposure, either for a single contaminant or a mixture. The exposure estimate is then used in an exposure–response model to estimate the probable magnitude of effects (like the “forward” use of distribution models in Panel B or C in Fig. 14.3). When using SSDs in this way, assessors must pay particular attention to ensuring that the exposure models or measurements reflect conditions in the field and are equivalent to the exposure metrics in the exposure-response models.

The outcomes of Criterion Risk Assessments and Conventional Risk Assessments can result in *follow-up assessments* (Suter and Cormier 2008):

3. *Management Assessments*. These assessments predict the results of one or more management alternatives that can be undertaken to reduce risks, like emission reduction, or the application of various remediation techniques. They consider risks, costs, benefits and legal requirements.
4. *Outcome Assessments.* These assessments can be used to determine whether the management action was effective (e.g., when concentrations have been reduced to the desired levels) and whether environmental goals have been achieved (e.g., when biotic communities have recovered).

In Sections 14.13 and 14.14, each of these assessment types is illustrated, with SSDs as a key part of the evaluation approach. It is shown that SSD-based results, with appropriate Exposure Assessment modeling, are specifically relevant for ranking ecological hazards, and that they can therefore be used to explore the results of alternative soil management scenarios. The theory and practical examples presented in this chapter show that *SSDs may help to improve effectiveness in the management or remediation of contaminated soils*.

14.4 Theoretical Basis of SSD Modeling

Though Fig. 14.3 is the illustrative basis for distribution-based modeling, there is more to say on the theoretical basis of distribution-based modeling in Ecological Risk Assessment. Hence, this section describes a range of technical aspects of SSDs, from basic principles to derivation and interpretation of SSDs and on refined issues like presenting confidence intervals and accounting for natural background concentrations of contaminants. We mainly focus on concepts, understanding and illustrations rather than on mathematical formulae; the latter can be found elsewhere (e.g., Posthuma et al. 2002b).

14.4.1 Why SSDs Fit the Risk Assessment Paradigm and Practices

SSDs are well suited to the general concept of Risk Assessment. Risk concerns both the probability of an adverse effect and the magnitude of such an effect. How distributional statistics relate to Ecological Risk Assessment is shown in Fig. 14.4.

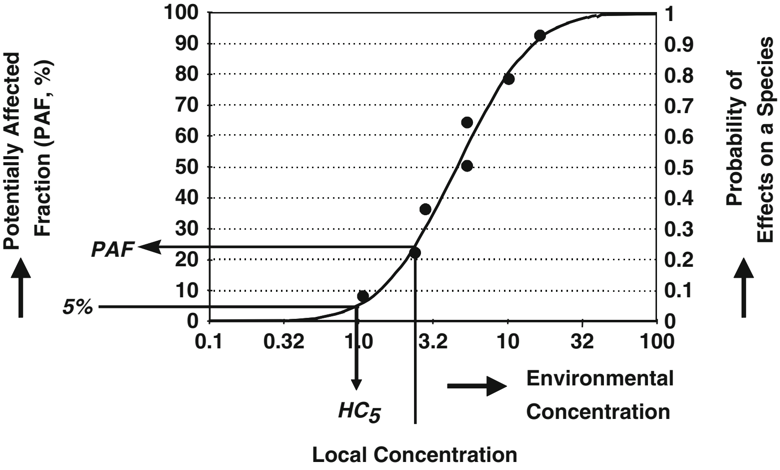


Fig. 14.4 SSDs and FSDs are suitable for Risk Assessment, since they address the two key elements of risk: probability of adverse effects, and their magnitude. This relates to the dual interpretations of the model. Left hand Y-axis: an SSD describes the relation between soil concentrations and the Potentially Affected Fraction of species or functions in an ecosystem. As the soil concentration increases, the PAF increases according to the shape and position of the SSD. Right hand Y-axis: an SSD describes the relation between soil concentrations and the probability that a randomly selected species from the test set would be affected when exposed to a soil concentration. The "forward" and "reverse" uses of SSDs in Risk Assessment are illustrated by the two arrows (see text)

14.4.1.1 Hazardous Concentrations

When Criterion Risk Assessment is applied to soil protection, the intent is to assure that the probability and the magnitude of effects is negligible or low. This can be achieved by setting soil quality standards, which are easy to apply in the daily practice of soil appraisal. If a soil sample does not achieve the standard, measures can be taken to reduce emissions or remediate the site.

In such cases, the SSD approach is used by selecting a maximum acceptable potential effect (Y-axis), and then reading off an "acceptable soil concentration" on the X-axis (Fig. 14.4, Y → X), which is sometimes referred to as "safe". This yields the aforementioned, *Hazardous Concentrations* (HC-values), like the HC5. The HC5 (a soil concentration) thus relates to a chosen risk limit (an effect measure). The HC5 (when derived from an SSD made from NOECs) is *the ambient concentration of a contaminant at which 5% of the tested species would be exposed at or above its No Observed Effect Concentration (NOEC) when raised in such a soil.* HC5-values for different compounds differ, suggesting differences in their toxicity to soil organisms. However, for all compounds, an exposure concentration at the level of the HC5 implies a similar level of predicted impact for the test species assemblage. SSDs for terrestrial Risk Assessments have been derived most often on the basis of total concentrations in test soils, and thus soil quality standards are mostly expressed on the basis of total concentrations.

When an HC5 based on NOECs is adopted as a regulatory soil quality standard, it can be understood that an exposure at- or lower than that HC5-concentration implies the protection of (at least) 95% of the tested species against adverse effects from contaminants in soil on those vital characteristics that are included in toxicity tests (like growth and reproduction). When the tested species sensitivity distribution resembles the natural sensitivity distribution, the HC5 represents a *95%-protection criterion*, which is an exposure level at which the structural properties of real soil ecosystems are almost all protected. Note specifically, that exceedance of the HC5-NOEC does *not in any way mean* an actual loss of 5% of species, although this sometimes pops up as implicit but wrong "gut feeling" about the meaning of the HC5 in discussions among stakeholders.

14.4.1.2 Hazard Potential or Toxic Pressure

When assessing *existing soil contamination* in Conventional Risk Assessments, both the probabilities of effects and their magnitudes increase with increasing soil concentrations, as in the field gradient in Fig. 14.2. SSDs closely link to this dual principle of assessing risks: essentially, *SSDs help to quantify probable impacts (assuming that the community is represented by the set of test species), given a soil concentration of a contaminant.*

In the "forward" use, SSDs can be applied to quantify the *Hazard Potential* (HP) of a contaminated soil sample, separately for each of the compounds present. In addition, by mixture modeling (see Section 14.10.6) a net impact of a mixture exposure can be quantified.

The term HP is conceptually consistent with the earlier introduced concept of HC and with the entity predicted (the fraction of test species that would potentially suffer at a given exposure). It is equivalent to the earlier term "toxic pressure" (e.g., Klepper et al. 1998), which suggests that a local mixture of contaminants would exert a quantifiable pressure on the tested biota when exposed in the soil, with higher probable impacts at higher toxic pressures. For the remainder of the chapter we will use toxic pressure instead of Hazard Potential. The extent to which these model results predict effects on real species assemblages in the field depends on the relationship between the SSD and the field species sensitivity distribution.

14.4.2 Extrapolation: From Probably to Potentially Affected Fraction

The outcomes of the approaches shown in Panel B and C of Fig. 14.3 would, in the case of the body height data, be relevant to the set of people attending the meeting. Likewise, in the case of SSDs, the predicted risks are relevant for the set of species tested.

Use of the distribution to predict how many people at *other* meetings could safely leave the room would however pose an *extra* problem, namely: a need to *extrapolate* the known distribution to another situation.

Analogous to this is the issue of *laboratory to field extrapolation* in Ecological Risk Assessment. This is a relevant concept because we usually do not (never?) have sensitivity data for species that occur at particular field sites, and field sites likely differ from each other and from the conditions at which the test data set was established. This holds both for the exposure conditions (and Exposure Assessment) and the local species (and their sensitivities). SSDs describe the distribution of the sensitivities of the test species, in their test conditions, *but not necessarily the distribution of species sensitivities and the exposure situation at a contaminated site*.

When test data distributions must be extrapolated, it is crucial to investigate how far a new situation could resemble the "test data set", that is the set of data to describe the distribution, and their exposure situations. For example, when the "test data set" of the workshop participants' heights are adults, it is easy to understand that a "blind" *statistical* use and interpretation of the model derived earlier would overestimate "risks" for children, but (far) less so for workshop attendants in general (since those tend to be populations of adults too). In a *practical* assessment, one could easily see what the output means (and what it does not) when considering the extra available data, and thus manage the case based on those extra data. For *any Ecological Risk Assessment, this means that one should scrutinize the resemblance of the contamination case to the "training data set", with respect to both the sensitivities of species and the exposure conditions.*

The issue of extrapolation has caused considerable debate (e.g., Hopkin (1993) versus Van Straalen (1993)). We resolve this issue at least semantically by recognizing a difference between the fraction of workshop participants or species that is *probably* affected and those at another workshop or in a natural species assemblage that are *potentially* affected. When the distribution of the training data set is used, the output of the distribution is a *Probability* that is directly valid for *the test set*. But after extrapolating to *another situation*, the extra uncertainty introduced by the extrapolation is frequently "captured" in a new term on the Y-axis, i.e., *potential*, in the phrase *Potentially Affected Fraction* (PAF).

If differences between the laboratory data set and the field are known, correction for those differences should be considered. There are various ways to correct exposure data for differences in conditions (see e.g., De Zwart et al. 2008b). These corrections often address bioavailability, which usually implies that estimated PAF-values are lower than when total contaminant concentrations are used. For effect data, various authors have discussed the extrapolation between life stages, test durations, and levels of organization (Solomon et al. 2008). However, laboratory to field extrapolations are, in general, highly uncertain.

As a useful alternative, it is becoming common practice to state at least qualitatively whether the field situation is likely to be more or less affected than modeled (as is possible in the adults-to-children issue in the example). Thus, a qualitative type of assessment output has been recently promoted as a common-sense solution to the extrapolation problem in various contexts; see e.g. EUFRAM (2006), regarding the evaluation of risk of plant protection products, and Risbey and Kandlikar (2007), on a formal scheme to handle Risk Assessment uncertainty.

14.4.3 The Conceptual Interpretation of SSDs: PAF and PES

As shown in Fig. 14.4, SSDs can be presented with two equivalent Y-axes. The statistical interpretation is the same, but the conceptual interpretations differ.

The interpretation of distribution-based Risk Assessments is best explained by a hypothetical experiment. Assume one has two contaminated samples, for which the Hazard Potential is estimated by Potentially Affected Fractions of 20 and 80%, calculated with SSDs constructed from EC50-values. This would imply that exposure of all tested species together in such soils would imply an effect level of 50% for probably 20 and 80% of these tested species. When assuming that the distribution of contaminated site species sensitivities resembles the calculated distribution, the same PAF-values (20 and 80%) are expected for the local species. Note that an ecologist would suggest in the latter case that 20% and especially 80% are underestimates of actual effects, due to possible indirect effects (such as in Fig. 14.2). For example, the loss of prey species would indirectly affect predator species, even if they could withstand the intoxication fully.

We do not know *which* species would be affected nor whether there are additional indirect effects. However, we have certainly *learned something about the potency of the contaminated soil to affect test species* and hence *we can interpret the "hazard" of this soil by extrapolation as a "risk level" for local species too, in the framework of practical soil management.* In other words: soils can be ranked according to their potential to affect test species (and possibly local species).

> SSDs re-calculate contaminant concentrations in soils (from the realm of chemistry, expressed in mg/kg$_{\mathrm{dw}}$ of a contaminant) into a toxic potency (in the realm of biology, in a dimensionless but conceptually meaningful unit, fraction of species), focusing on the fractions of species likely to be affected. PAF is, in fact, a characteristic of a soil to be hazardous to its inhabitants. It is *not* an ecological entity.

In the same hypothetical experiment one can also interpret the SSD according to the concept of *Probability of Effects on a Species* (Fig. 14.4). In this approach, the hypothetical experiment proceeds as follows. Imagine the same soils, but now the assessors are concerned with risks to a particular species, and that species may be considered to be a random draw from the set of all species, which is estimated by the distribution of the set of tested species. In this case, the probability that this species would be affected by the contamination would be 0.2 and 0.8, on a probability scale of 0 to 1. Again, *this interpretation tells us something about the potential of a contaminated soil to affect a randomly chosen species that could reside in that particular soil.*

The *PES* and *PAF* interpretations can be used interchangeably, but PAF is generally used in practice as the unit of toxic pressure.

14.4.4 Discussions of SSDs, Assumptions and Interpretation

Since SSDs are statistical models based on minimal ecological knowledge, they have been discussed intensely since their invention, amongst scientists and with regulators. The discussions have been summarized extensively in Posthuma et al. (2002b). The discussions have focused primarily on the "reverse" use of SSDs, the derivation of quality standards.

The issues as discussed can be divided into two different types:

- technical issues;
- validation issues.

Technical issues include the selection and quality assessment of test data, handling multiple test data for a single species, the choice of the statistical approach (including non-parametric methods and fitting of a standard parametric model), the choice of a risk limit (like 5%) to derive standards, the choice of the estimated HCp or its lower confidence interval to be used in regulation, the choice of test endpoints (e.g., LC50s or NOECs), the required number of input data and the inclusion of different taxonomic groups. The main driver for these discussions was to get the "best and most standardized SSD-approach" which should be optimally representative for "the" soil system. "The" is placed in quotes, because there is no uniform soil system. Technical discussions of SSDs have paid surprisingly little attention to the diversity of soil systems *versus* the wish for standardization. Given the emphasis on the reverse (standards) use of SSDs, validation discussions mainly focused on the question whether the standards derived with SSDs were indeed sufficiently protective, and whether any extrapolation of laboratory test data to appraise ecological impacts in the field can be valid.

These discussions have been worthwhile in a scientific sense, since they have led to improvements in the methods. However, since they often address a standard method (e.g., on the number of required test species data, and on representation of different taxonomic groups as required for standard setting), one should take into account that a standardized method relates strictly to the idea of a single problem definition. An example is the derivation of soil quality standards for a compound in a country. Such standards often have a legal status, and this implies that they should be consistent in their derivation. This has led to the regulatory choice of well-defined uniform methods, and the fixation of those methods in guidance documents. When the problem definition, however, requires a "forward" use of SSDs (i.e., Conventional Risk Assessment), the uniform choices made for the derivation of quality standards need not the best ones. For example, when assessing the net risk to the local soil biota caused by a local metal mixture, one is able to apply scientific knowledge *not* used in the standard SSD-method. For example, there is clear evidence that the effects of metals in acid soils are much higher than in neutral soils (Janssen et al. 1997). Hence, this should show up in the site-specific (forward) Risk Assessment, such that a certain metal concentration in an acid soil is indeed more risky than one in a neutral soil. This can be solved, for example, by selecting ecotoxicity input data for forward-SSD use from tests in relatively acid test soils, to implicitly address the high availability of metals in these tests. Re-use of the standard choice methods (e.g., taking the average of all available data, irrespective of pH) would result in risk levels indifferent to soil acidity, which is wrong (see e.g. Cleven et al. (1993) for some tabulated examples of large pH effects; page 60 in the cited article). In other words: the uniform methods adopted for the "reverse" use (like averaging all input data for a species) need not be applicable to problem solving in the "forward" use (selecting appropriate data).

In general, we think, three rules should be applied when discussing technical issues and choosing technical approaches in SSD modeling. First, the specific problem definition of an environmental case should be the guide to appropriate choices in the Risk Assessment with SSDs. Second, when a statistical model is used, it should show a sufficient fit to the data. Third, when a problem occurs frequently (such as in setting quality standards for chemicals), fixation of SSD methods may be efficient and necessary. The latter is especially the case for legally binding SSD output. Evidently, all efforts should focus on results that are valid for the problem and situation.

14.5 Validity of SSD-Based Output in Ecological Risk Assessment

SSDs are statistical models derived from laboratory toxicity data that should be applied in combination with appropriate exposure models and extrapolations, to obtain potentially meaningful and useful output. The big question is: how can these statistical models predict anything relevant for ecosystems?

This practical question has triggered validation studies. Suter (2002) has identified two levels of validation of models in Ecological Risk Assessment:

1. validation by practical use;
2. validation by model-derived values resembling the relevant field phenomena.

To begin with the former, various organizations have formally adopted SSDs as a method in Ecological Risk Assessment (for an overview see Posthuma et al. 2002b), and especially for the derivation of soil quality standards. When soil quality standards are used, they are continuously being challenged in practice. Cases where impacts are shown when they are unexpected and cases where standards are unrealistically low will lead to public and policy debate, and eventually to development and adoption of more defensible methods. Applications of the SSD-based U.S. Ambient Water Quality Criteria, for example, have been legally challenged over the years since they were published in 1985. So far, they have withstood those challenges, like the soil and water quality standards in Europe (which in part relate to the use of SSDs). Nonetheless, there are ongoing debates on certain contaminants (such as zinc) and certain system issues, such as accounting for natural background concentrations. Such discussions result in technical and conceptual improvements in practice.

The latter type of validation implies scientific comparison between predicted and observed impacts. Two tracks can be recognized here: studies that focus on point estimates (such as HC5) and impacts at such exposures and those that focus on complete SSD curves in relation to changes in species assemblages in the field.

The first track is a phenomenological study on soil quality standards in relation to observed occurrence or absence of effects in contaminated field sites. When considering Fig. 14.2, the onset of responses might be compared to the position of the SSD-predicted HC5, to check whether the HC5 is indeed a “safe environmental concentration”. In various studies, the HC5 based on NOECs (and often assumed to represent a concentration at which ecosystem structure is protected) appeared to be positioned in the “non-response left tail” of biotic response curves. Note that, by this criterion, any way to obtain a low estimate of a soil quality standard would also yield a valid result (a soil quality standard that is higher than the ecosystem no-effect level and thus sufficiently protective). For example, dividing the lowest NOEC by an uncertainty factor of 10 or 100 (often an alternative approach to derive standards) is also valid in this sense. As noted below (Section 14.13.3) it is of practical importance to derive standards that are *not* unnecessarily low (e.g., by choosing a very high safety factor or the lower confidence limit of a chosen HCp-level), since that would imply (public) concerns where they are not realistic, and possibly extremely low cost effectiveness of Risk Management.

In some legislation, there is also a trigger value for considering remediation, and that trigger is derived from an SSD. In the Netherlands, for example, the HC50 of an SSD-NOEC model is used to select cases requiring further study, as a step-up to potential remediation. Similar to the HC5-validation approach, predicted HC50s (the concentration at which 50% of the species would be exposed beyond their NOEC) appear to be positioned either in the neutral-response part of the response curve of field assemblages, or in the steep down-going part of this curve, but (so far) never in the right tail where all biota suffer (Posthuma et al. 1998). In other words: such HC50s appear to signal the presence (or near-presence) of adverse ecological impacts in the field.

Examples of this kind of study are summarized in Posthuma et al. (2002b). More recent studies have signaled some observed field responses at the level of an HC5 (see DEFRA (2005) and Frampton et al. (2006)). In particular, effects have been found in field tests of pesticides. But note here the possibly specific impact of the clear exposure scenarios (short, peak exposures after deliberate use) and the possible focus and intensity of the effect studies that can be made at such moments.

The second track of validation studies has been recently developed, and considers the association between predicted impacts and observed species loss over the whole concentration range of contaminants in the field, based on very large biomonitoring data sets. Because of the occurrence of mixtures in such databases, the SSD-based predictions also incorporate probable mixture impacts (the toxic pressure now concerns the multi-substance PAF, msPAF; see Section 14.10.6). Figure 14.5 presents two studied cases on the degree of association between predicted (X, acute toxic pressure, based on SSD-EC50s) and observed species loss for fish species in Ohio (U.S.) rivers (De Zwart et al. 2006; Posthuma and De Zwart 2006), and for freshwater invertebrate species in England and Wales (De Zwart et al. 2008a). Given the results of both studies, five things are noteworthy:

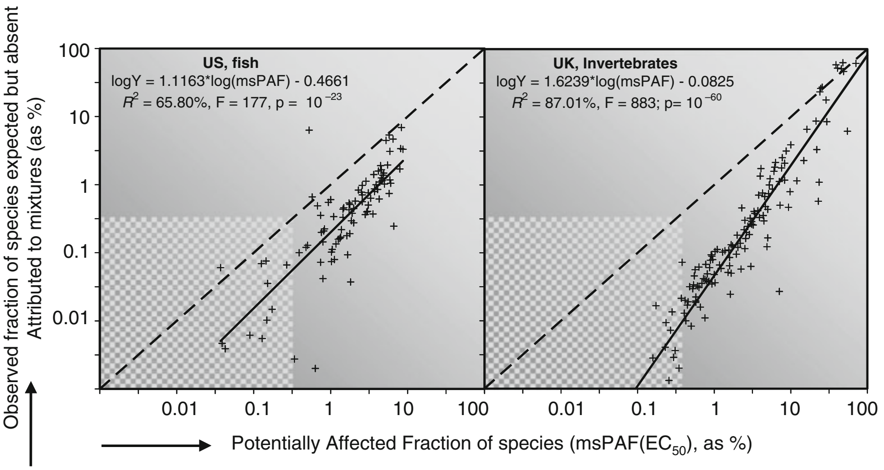


Fig. 14.5 Association between (SSD-) predicted (X-axis, msPAF) and observed fraction of species loss attributed to those mixtures (Y-axis), derived from diagnostic modeling on (bio)monitoring data concerning fish species in Ohio (U.S.) rivers, and freshwater invertebrate species in England and Wales. The toxic pressure data are based on SSD-output for various contaminants, yielding an msPAF based on EC50 data for each site. The Y-data are based on a diagnostic method, separating out the probable contribution of actual contaminant mixtures from other stressors as causes of species loss (De Zwart et al. 2006). Data points refer to XY-consolidated averages. All the sites with (at the first decimal level of their logs) comparable values are represented by one single cross. The dotted line represents the ideal hypothetical condition, where predicted fraction of species affected equals observed species' loss attributed to toxicant mixtures. The dotted grey areas indicate that there is a range where both X and Y values are lower than realistic values (any species' loss can *per se* not be lower than one, regardless of the chosen taxonomic resolution). The solid lines were fitted to the observations and were highly significant. Data from De Zwart and Posthuma (2006) and De Zwart et al. (2008a)

1. There is an association between predicted (X) and observed species loss (Y), with a highly positive slope: apparently increased acute toxic pressures imply increased probability of species loss in the field.
2. The predicted species loss (X) is generally higher than the observed loss attributed to mixtures (Y), indicating species loss due to other stressors and/or an association between EC50 and species loss that is generally not 1:1, apart from the highest levels of mixture exposure.
3. At the highest toxic pressure levels, the predicted and observed values tend to be similar, at least for the invertebrates.
4. There is huge variability (more or fewer species may be lost than predicted) probably due to natural variability and other stressors.
5. There are no observations where there is low toxic pressure (X) but an associated high species loss (no data upper left corner), nor is there a high predicted loss and no observed loss (no data lower right corner).

We conclude from the scientific validation studies, that

- an HC5 that is estimated from NOEC values (or similar measures of none or little effects in test systems) is an estimate that for many known cases indicates the absence of no or very limited adverse effects on policy-valued ecosystem characteristics (like biodiversity), so that such SSD-based standards can be used for deriving protective soil quality standards;
- the estimation of Potentially Affected Fractions, especially for SSDs based on EC50 values, provides insight into the potential loss of species at contaminated sites that permits the relatively robust ranking of contaminated sites from low to high probability of toxic impacts, but levels of effects need not be well predicted because many other stressors contribute to causing species loss.

We explicitly state that we have rather few examples, and especially few for soils and sediments in comparison to examples for aquatic systems, to support the generality of these conclusions. Nonetheless, the incidental studies seem to imply that increases in SSD-predicted risk levels are related to increased impacts in natural systems. An SSD-based risk prediction may be not robust in the sense that such an estimate exactly predicts *what* will happen to local species (neither which one will be affected, nor to what extent), but *that* large differences in (ms)PAF certainly seem to imply large differences in potential for species impacts or species loss. This phenomenon implies that SSDs are useful at least for risk ranking.

14.6 SSDs and Ranking of Contaminants or Sites

14.6.1 SSDs and Ranking Contaminants

The two ways of using SSDs in Ecological Risk Assessment were already introduced in Fig. 14.4. When these two approaches are used to assess contaminants or contaminated sites, and in view of the validation studies, the key interpretation of the SSD output is: *ranking*. Note that this is not peculiar to SSDs. Relative ranking is more reliable than absolute estimation for all assessment techniques.

When adopting standard guidance protocols, one can derive Hazardous Concentrations for *p* percent of the species for all possible contaminants being tested using the Y→X approach. When this is done for two or more contaminants, two or more SSDs are generated, and two or more HC5 values result. The output of this conceptually implies an *ecotoxicological "ranking of contaminants", in the sense that the lower the HCp-value, the more toxic the contaminant is for the tested sets of soil organisms.*

Some issues arise when using SSDs to rank contaminants in soils. First, the ranking can be based on the HC5 values, showing which contaminant is the most toxic at this exposure level and how the other compounds rank. But when another HCp-level is used, rank orders might change as a consequence of SSD-slope differences. Second, there is an option to trigger less debate on the ecological meaning of

SSD-modeling and HCp, by focusing specifically on this ranking interpretation. When focusing on HC5s for ranking compounds, one benchmark contaminant (for example: cadmium) can be chosen, and all other compounds can be characterized by a statement on how many times more or less toxic another contaminant is for (tested) soil organisms relative to cadmium. That is: if the HC5 of cadmium is 5 mg/kg$_{\mathrm{dw}}$ soil and that of contaminant X it is 15 mg/kg$_{\mathrm{dw}}$, then the contaminant is estimated to be three times less toxic. An even better way of such scaling could be provided when having two benchmarks (such as on the scale of degrees Centigrade for temperature, freezing and boiling of water). Such a ranking in itself tells us something about the relative environmental burden of a contaminant when released in the environment, as compared to the benchmark(s) (all relative to the tested sets of species).

The ranking of the relative environmental burden of a set of contaminanats according to their HCp is illustrated in Fig. 14.6, and is shown to illustrate the robustness of this way of looking at HCp-estimates. Two independently collected datasets of tested contaminants were chosen from literature. The sets were used to obtain two HCp-estimates for each of the contaminants according to independent protocols. The two assessments were done by Jänsch et al. (2007), using EC50s as input data, and Swartjes (1999), using NOECs as input data. Hence, the assessments differed by the choice of input data in a major way. The figure shows that both assessments resulted in a similar rank order of the HC5-NOECs and HC5-EC50s for soil organisms. There is also some variability (distances of point estimates from the fitted rank order lines). As a consequence of the different input data, the lines differ from an equality line (slope 1, intercept 0). The HC5-NOECs of Swartjes are generally lower than the HC5-EC50s from Jänsch et al. The findings imply that the

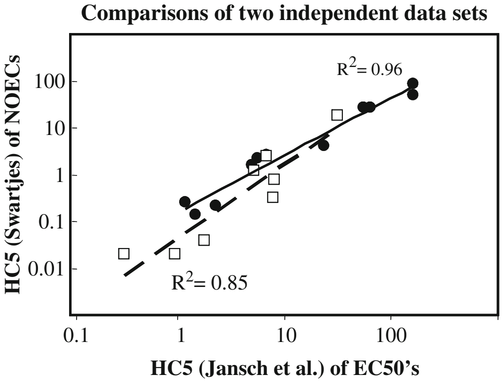


Fig. 14.6 Comparison of HC-values for a set of metals (*black circles*) and organic contaminants (*white squares*) from SSDs for soil organisms. The fitted regression lines (continuous and broken, respectively) suggest a high similarity in ranking results between both methods. Deviations of the markers from the lines identify statistical issues, like different choices of data selection criteria and model choices. Data from Swartjes (1999) pertain to HC5-NOECs and of Jänsch et al. (2007) pertain to HC5-EC50s. Therefore, the estimates generally differ by approximately one order of magnitude (in accordance with an often-found acute-to-chronic ratio of 10)

interpretation of risk ranking across compounds seems to be very consistent and possibly more appropriate and precise than the interpretation of quantification of ecological risks at the level of HC5-NOEC in terms of impacts (as treated in the validation study section). The ranking of contaminants with respect to their potential environmental risks appears rather robust.

14.6.2 SSDs and Ranking Sites

For the other use, the X→Y read-off, one can assess the potential of contaminated soils to affect test species assemblages (PAF) or the probability that a randomly selected test species is affected (PES). If one would consider a suite of contaminated sites (such as a suite of contaminants), one can predict the fractions of test species that would be affected in those soils. When re-thinking the conceptual meaning of the output as obtained, the output in fact provides an *ecotoxicological "ranking of the contaminated soils (based on toxic pressure quantification, through PAFs)", in the sense that the higher the estimated PAF or PES, the more toxic a soil is for tested soil organisms (Klepper et al.* 1998), *and probably also for natural species assemblages when exposed in such a soil.* Again, one could think of one or two contaminated sites being the benchmarks for comparative ranking purposes ("this site is three times more hazardous than benchmark site X"), but there may be practical problems in selecting such benchmarks. In the Section on validation, Fig. 14.5 showed how *sites* are ranked regarding the relative expected impacts of site contamination, like the contaminant ranking of Fig. 14.6.

14.6.3 SSDs, Rankings and Weighting in SSDs

In ranking compounds or sites, one should acknowledge here that one treats all species as equal points in SSD models. Some authors have proposed weighted SSD modeling to account for differences in ecological or societal importance of species (see Section 14.8.2). The outcome of the simple ranking with SSDs might change if valued species are weighted. In other words, a ranking of sites with unweighted SSDs ranks the array of sites according to only their potential to affect the species in the data set of tested species (the Hazard Potential interpretation *sensu stricto*).

14.7 SSDs and Cost Effectiveness of Environmental Management

Environmental management is often expensive.

Preventing soil contamination using Criterion Assessments may imply major costs, e.g. when the environmental impacts of newly produced contaminants (like pesticides) are high. Compounds might be banned when expected risks are too high, while use limitations may in some cases be prescribed to keep soil concentrations below the soil quality standard. Banning a pesticide implies that the multi-year development process of the contaminant does not bring profit to the producer, and whether this is based on a right or wrong model, there are considerable costs involved.

Remediation of contaminated sites may be very costly. If a policy framework prescribes specific actions when certain triggers have been exceeded (e.g., an ecologically based trigger-concentration for remediation), a government wants to be sure about the fact that (a) there is real hazard, and (b) there is sufficient budget, as well as (c) that the budget is spent best, with maximum ecological improvements as a result.

Instead of prescribing certain actions when soil quality standards are exceeded, this book advocates, in Chapter 1 and Fig. 14.1, tiered Risk Assessment schemes. We thus propose that, when a soil quality standard is exceeded, there is a second option for practitioners, namely: conventional Risk Assessment of how serious the exceedance is in terms of local or future soil use. Factors such as differences in sorption or compound breakdown may be addressed in the higher-tier approach, with SSD-modeling, or other (even higher-tier) models. Relatively labor-, time- and budget-intensive methods such as tests of site media or from field impact studies are not justified, unless lower tier studies demonstrate a need. Exploration of possible impacts using SSD-models is cheap and quick, provided that the basic input data are available.

(Cost) effectiveness is an important practical driver in the context of decision making processes, and we present some examples (below) in which we illustrate that SSD-based risk modeling of alternative management scenario's can be helpful. The examples of SSD use are practical illustrations of potentially large differences in cost effectiveness of alternative Risk Management scenarios.

14.8 Practical Basis of SSD Modeling

Now that it is clear what SSDs mean (and don't mean) we turn to *the construction of* SSDs. Practical application not only depends on meaningful output and accounting of extrapolation problems, but also on ease of use. For SSDs, the situation is simple. SSDs can be easily made with only two ingredients:

1. input data, which are ecotoxicity test data for a particular contaminant and various species or functional traits, and
2. an appropriate data handling approach.

14.8.1 Ingredient 1: The Input Data

14.8.1.1 Raw Input Data

The input data for an SSD can be collected by original research, from the literature or from databases in which such data are stored.

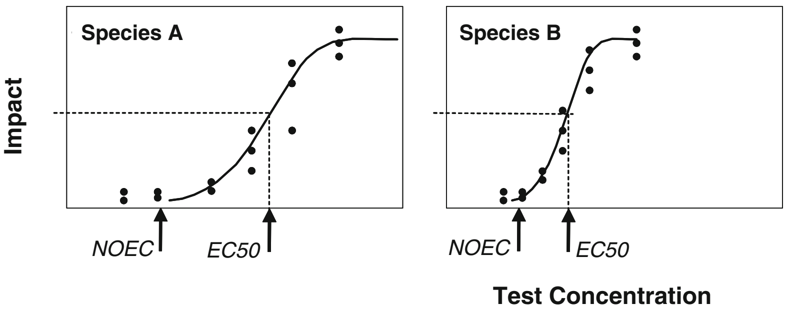


Fig. 14.7 Plotted laboratory toxicity data, collected for two species (**a** and **b**), show the principles of designing toxicity tests and collecting summary data on species sensitivities. (**a**) is a relatively insensitive species (position on the X axis) with a low increment of response at increases of exposure (slope), and (**b**) is more sensitive species with a high increment of effects at increasing exposure

Input data consist of sensitivity data, collected usually in ecotoxicity tests of various species. To obtain such data, species are usually exposed to multiple concentrations of the contaminant of interest, under artificial test conditions (e.g., artificial soil, or a well-known test soil under controlled conditions). After a period of exposure, either the survival of the exposed organisms is scored, or the performance of vital characteristics is quantified (e.g., with parameters describing reproduction or growth performance).

Summary sensitivity data are derived from such tests, per species. Collected data may be critical (chosen) concentrations, like *No Observed Effects Concentrations* (NOECs, the highest concentration without significant effects), *Median Effective Concentrations (e.g.,* EC50, the test concentration that affects 50% of the test population), or other test endpoints. The principles of designing toxicity tests and collecting summary data on species sensitivities are shown in Fig. 14.7.

Original tests may be necessary, for example, when data are lacking for contaminants for which regulatory action is deemed to be needed (e.g., due to frequent use or occurrence of the contaminant at contaminated sites).

14.8.1.2 Pre-Treatment of Input Data

Pre-treatment of input data can include practices such as sub-selections based on quality assurance criteria (e.g., sub-selection of data collected under Good Laboratory Practice), selection of data with consistent exposure durations, or conversion of data to standard units, or calculating the median sensitivity when more than one effect data value is present for a specific species and endpoint. When SSDs are formally adopted in regulations, it is common that there are *Guidance Documents* stating how data pre-treatment should proceed. For example, in the U.S. EPA's guidance for water quality criteria, multiple data for a species are combined by taking the geometric mean. When there is no prescribed guidance, data pre-treatment should match the context of the case.

14.8.1.3 Example Data Bases

Various data bases have been made to facilitate the collection of input data for SSDs. Examples of compilations of all kinds of ecotoxicity data are the U.S. EPA's *Ecotox database* (U.S.EPA 2002) and the Dutch RIVM *e-toxBase* (Wintersen et al. 2004). Many researchers have their own data collections.

14.8.2 Ingredient 2: The Statistical Approach

14.8.2.1 Options for Model Choice

Figure 14.3 shows that one can use raw sensitivity data directly (bar diagram), and that one can use models, like parametric curve fitting, or more complex approaches, like nonparametric or Bayesian regression.

Frequently, SSDs have been derived using a log-logistic or a log-normal model. Alternatively, multi-modality of the SSD can be considered when the bell-shaped curve has two peaks, see e.g., Aldenberg and Jaworska (1999). In such cases, there are *two* subsets of ecotoxicity data. A specifically sensitive subgroup can be distinguished, such as insects in the case of insecticides. Non-insects can then be modeled with a separate Normal distribution shifted to the insensitive side. In this case, the user may want to make two SSDs, one for each subgroup of species, instead of one bimodal model.

The user may wish to give different weights to the different input data (instead of weighting all data equally; see, e.g., Duboudin et al. (2004)), because some of the species may be considered more important or relevant for the soil health. Earthworms, for example, are sometimes considered to be "ecosystem engineers" (see, e.g., Jouquet et al. (2006)), due to their large capacities in soil turnover and aeration, and arguably they might be given a higher "weight" in deriving the SSD for a contaminant. How to give those weights is an issue of expert judgment, and is not specific to SSD-modeling.

14.8.2.2 Selecting a Model

All statistical approaches have advantages and disadvantages. Not all of these issues are treated here. Posthuma et al. (2002b) provide references to the different models and approaches. For practical users it is *not* necessary to fully understand the statistical refinements of each of the methods, although all mathematical approaches and practice details differ in kind. Three theoretical issues should be considered in the model choice:

1. the fit of the chosen model to the data (misfits may identify specific aspects of the input data);
2. despite an overall good fit, the possible bias of the model in the concentration range of interest, and finally
3. the validity of the output for the case of concern

14.8.2.3 Software

SSDs can be derived using any software package with which one can fit distributions, and there are many. Examples of freely available software specifically for SSDs are *ETX* (Van Vlaardingen et al. 2004), *MS-Excel spreadsheets* for Weighted SSDs (Duboudin et al. 2004), an executable named *OMEGA* (Beek et al. 2002), and web-based platforms like www.risicotoolboxbodem.nl and WEBFRAM-SSD modeling (http://defrarisk.cadmusweb.com/Generic1.aspx). Both ETX and the WEBFRAM SSD software run three Goodness-of-Fit tests (Anderson–Darling, Cramer–Von Mises, and Kolmogorov–Smirnov), which address the first issue by focusing on the over-all fit for the whole range of input data.

14.9 Statistical Issues in SSD Modeling and Interpretation

So far, we have introduced the concept of distribution-based modeling, two conceptual ways to use SSDs in Ecological Risk Assessment, software tools, and validation. In practice, SSD-modeling also requires a framework within which results are collected and interpreted. In this Section, we address statistical issues in SSD modeling and interpretation. The next Section discusses other issues, for example related to exposure conditions and to handling mixtures.

14.9.1 Minimum Data Numbers and (Mis)Fit

Many authors have discussed the minimum amount of data needed to do SSD-modeling. Some agencies have adopted specific minimum numbers of species while others derive SSD-output based on cases with few input data.

Some regulatory frameworks formally require a minimum data set for standard setting. An example of the *Guidance* on the use of SSDs for deriving standards for the EU is as follows: "*Confidence can be associated with a PNEC* [i.e., the Predicted No Effect Concentration, like an HC5] *derived by statistical extrapolation if the database contains at least 10 NOECs (preferably more than 15) for different species covering at least eight taxonomic groups. Deviations from these recommendations can be made, on a case-by-case basis, through consideration of sensitive endpoints, sensitive species, mode of toxic action and/or knowledge from structure-activity considerations*" (European Communities 2003). In the U.S.A., data from eight different genera of fish and invertebrates are required for setting National Water Quality Standards (Stephan et al. 1985).

We do not discuss the justification for these chosen values, because they are part of specific regulatory soil protection contexts, and they are for the derivation of generic legal standards (Y→X), not necessarily for the more specific use of SSDs (X→Y).

As an alternative set of criteria for cases when there is not (yet) a regulatory fixation of methods, we propose a two-way approach to help in deciding whether an SSD-model is appropriate for a problem definition. First, in a statistics-based approach, one can check whether the model fits the data (or not), by applying goodness-of-fit testing, and set *a priori* fit criteria to accept or reject a fitted model. A software program like ETX supports this kind of statistical evaluation (see Van Vlaardingen et al. 2004). Models that show a bad fit to the data can thus be rejected on a chosen statistical criterion, irrespective of the number of data and the set of test species.

Second, the input data should be selected in view of the problem definition. For example, for judging soils at low pH, one could select all test data with a pH near the pH of the studied soil.

It should be noted that rejecting input data by any data selection criterion has a trade-off. Low data numbers imply wider confidence intervals (see e.g. Fig. 14.3), and sometimes rejection of an SSD model, implying that another model needs to be chosen. The past selection of NOECs for the derivation of soil quality standards is, in terms of this tradeoff, a surprising one. Usually there is much more acute toxicity data, so that sensitivity distributions can be derived with less (statistical) uncertainty based on such data. If acute data had been chosen for setting protective standards (such as that explored by Kooijman (1987)) this would have required *post hoc* acute to chronic extrapolation. The acceptability of a particular criterion for data selection for an assessment depends on the case.

14.9.2 Presenting Confidence Intervals

SSDs, like many statistical models (with large or small sets of input data) can produce both point estimates and confidence intervals of those estimates. Confidence intervals represent which values the "true" value could have, given the variability of the available input data. Beware, that this only pertains to the *statistical* meaning of confidence (not to extrapolation in the case of ignorance of e.g. the sensitivities of the field species).

An example of a statistical analysis of confidence intervals in SSD-outputs is provided in Fig. 14.8, for the statistical confidence interval of an HC5. When the SSD-output is used for the derivation of a formal soil quality standard, the statistical analysis is followed by a policy choice, because a standard cannot be uncertain; a formal soil quality standard cannot have the form "a maximum tolerable concentration of 5 plus or minus 2.5 mg/kg_{dw} soil". Hence, the regulatory context requires a choice, either for the *p* in HC*p*, or (when more precaution is taken) for a lower confidence bound on the HC*p* (e.g., resulting in a value of 5 minus $2.5 = 2.5\,\text{mg/kg}_{\text{dw}}$ in the example shown above).

In a similar way, one can quantify the confidence interval for estimates of PAF (Y), given an actual soil concentration (X). In this case, it is possible to present a comparison between two contaminated sites, by demonstrating for example that site A would induce impacts for 50% of the species (confidence interval 45–55%), and site B for 80% of the species (75–85%). Given those small and non-overlapping confidence intervals, site B clearly presents a greater risk for the test set (and probably also for field species).

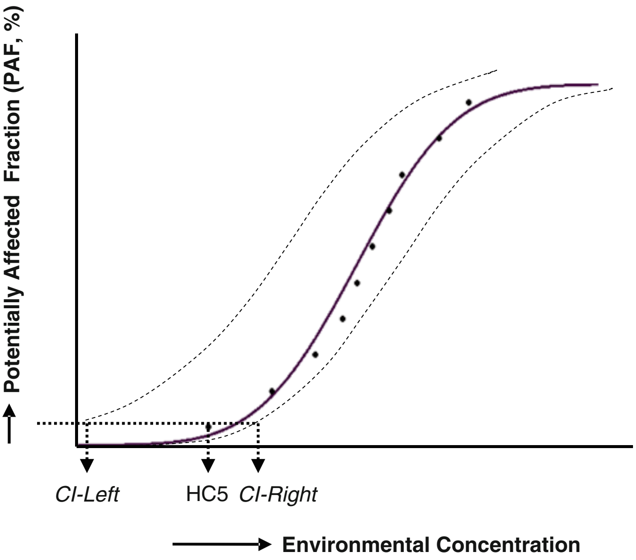


Fig. 14.8 Presentation of statistical uncertainties of SSD results. The outer curves represent the 95% Confidence Intervals (CI) of the fitted SSD (the central curve). They are used to estimate confidence limits of the HC*p* values (shown here: HC5). The dots are the original data and the fine line is an empirical step function

14.9.3 Interpreting Statistical Confidence Intervals

If one fits various models to the same data set, multiple models often show *a similar degree of fit to the data.* That is: all S-shaped mathematical functions fit well on a (log) concentration scale, or none of them fit. What is the meaning of this, given the possibility to present and interpret point estimates of HCp or confidence intervals?

When various models show a similar degree of fit (and overlapping confidence intervals) the SSD-output (in the format of estimated values of e.g. an "acceptable concentration" like the HC5) can nonetheless be numerically dissimilar, especially when the required output is positioned in the tails of the distribution, where small changes in estimated response are associated with relatively large differences in soil concentration and where confidence bounds are wider.

This apparent contradiction (similar model fit, different outcomes) has caused various intense debates at the interface of science and policy. *Statistically* indistinguishable differences in the numerical value of the HC5 (thus being scientifically without serious implications) may practically mean the difference between permitting and banning a pesticide or other newly developed chemicals. This problem is illustrated in Fig. 14.9. If one considers this graph in the context of the uncertainty intervals that can be estimated (see Fig. 14.8), it is clear that the three estimated risk limits (HC5) in Fig. 14.9 are *scientifically* estimates of the same "true" HC5 (that is: all three numerical HC5-estimates are positioned within the confidence intervals of all three models shown). For societal use, however, the three possible *numerical*

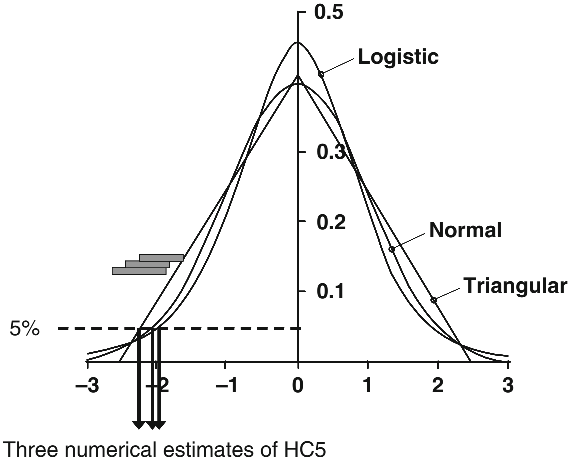


Fig. 14.9 Three different statistical models (logistic, normal and triangular) can fit data (not shown) equally well, but result in dissimilar estimates of numerical outcomes (in the example: the 5th percentile cut-off of impacts would result in three different numerical values of the HC5. Note that the X-axis represents a log-scale, so that the HC5-values are all near 0.01 mg/kg$_{\mathrm{dw}}$. The three horizontal bars indicate that the Confidence Intervals of all three HC5-estimates overlap

values of HC5 would lead to three different proposals for a soil quality standard, while only one could be published as a formally adopted value. The model choice thus always matters in terms of societal and practical impacts, but much less so statistically.

To give an impression of the possible practical impacts of choosing alternative statistical approaches for the same input data, as in Fig. 14.9, we refer to Verdonck et al. (2001). These authors concluded on the basis of statistical comparisons between alternative SSD-modeling approaches that those approaches vary in statistical robustness *per se*, that they differ in their sensitivity for decreasing sample sizes, and that the numerical estimates for HC5s as generated may differ by a factor of 5. Most importantly, they said there is no *a priori* reason to prefer any of the possible statistical approaches. These findings illustrate three relevant issues. First, the factor of 5 mentioned by Verdonck is much smaller than the uncertainty factors (of 10 and 100) that are often applied in Effect Assessments, according to various guidance documents, to derive a quality standard from (e.g.,) the lowest test NOEC. Such uncertainty factors are often 10, 100 or 1000. Second, though the point estimates may vary by a factor of 5, there may be overlap of the confidence intervals of the estimates generated by different methods (see Fig. 14.9). Third, based on validation studies, on the concentration scale of a field gradient *all* HC5s may be in the range of "no statistically significant or visible signal of response for the endpoint of concern" (compare neutral response concentration range in Fig. 14.2). Whether these three issues apply in other cases could be a subject of study, e.g., by fitting alternative models to the data, and by comparison of the results amongst each other and with results from an uncertainty factor approach. Next to these statistical explorations, it is generally undoable to find compound-related case studies for checking the validation issue.

Because a factor of 5 may make a significant difference in the regulation of contaminants (Y→X) and in the land area requiring remediation, it is important to use a defensible and consistent method to select a model. For dose-response modeling, the U.S. EPA for example routinely uses Akaike's information criterion (AIC) to select the best dose-effect model from a family of optional mathematical dose-response models (http://www.epa.gov/ncea/bmds/bmds_training/methodology/intro.htm). The AIC is an approach that compares several competing statistical models when fitted to the same data set, and determines the best model in terms of the information obtained from the data.

14.9.4 Options to Handle Small Sets of Input Data

For most new synthetic contaminants, no soil test values are available and for most others the numbers of data are not sufficient to estimate an SSD. The European inventory *EINECS* (European Inventory of Existing Commercial Chemical Substances) alone contains more than 100,000 contaminants (http://ecb.jrc.ec.europa.eu/esis/index.php?PGM=ein), while the RIVM e-toxBase contains data for approximately 5000 contaminants.

The real or expected lack of input data for SSD-modeling implies a necessity to consider methods for estimating test results. In such cases, existing ecotoxicity data are used to predict the ecotoxicity of compounds for non-tested species (or species groups) of interest, and to predict chronic input data based on available acute ecotoxicity data. De Zwart (2002) for example has derived "rules of thumb" for deriving the shape and position of SSDs of untested contaminants from test data for other chemicals. The procedures are based on patterns in shape or position of SSDs, which in turn relate to grouping of contaminants by Toxic Mode of Action. Aldenberg and Luttik (2002) have proposed specific methods to handle small toxicity data sets, again by using toxicity data from "similar" tested contaminants. Finally, the U.S. EPA provides inter and intra species extrapolation models for contaminants for which data are lacking. The ACE model addresses Acute-to-Chronic Extrapolation, and makes use of the fact that for many contaminants there is a wealth of data on acute effects, relative to chronic exposure test data (Ellersieck et al. 2003). It is possible to extrapolate from acute to chronic data due to existing data patterns, yielding the possibility to derive chronic SSDs when only acute data are available. The ICE model is concerned with Interspecies Correlation Extrapolation, and makes use of known correlations of toxicity values (e.g., EC50) between a surrogate species (frequently tested) and an untested species or taxon of interest (Mayer et al. 2004).

14.9.5 Handling the Possible Causes of Misfit

SSD models may adequately fit the data, or not. Various statistical tests and diagnostics are implemented in SSD-software to investigate model fits. In the case of a sufficient fit of the model to the data, the assessor may decide to report this, and use the model. The statistically sufficient fit, however, may *not* be interpreted as any kind of mechanistic justification for using the model: a statistical test may just signal misfit, in the case of which the outcomes of SSD-modeling might be practically doubted for that reason alone.

In the case of misfit, there is a scientific trigger to reconsider the data. Are there, for example, two subsets of species (a sensitive and a non-sensitive subset), as in the case of bimodal SSD for an insecticide, which was tested for a wide array of species? In such cases, the misfit might trigger deliberate choices of the use of data. Eventually, this may imply deriving an HC5-risk limit on the basis of "target species" only, or the assessment of risks for various subgroups of species. Examples of sensitivity differences between subgroups of species are provided by Frampton et al. (2006). In this case, a refined Risk Assessment with SSDs is possible (see Section 14.10.7).

Alternatively, the lack of fit may be due to one or more low quality data. In particular, the set of test data may contain tests in which the tested species are exposed in conditions that are unnatural and stressful for the test species (Jänsch et al. 2005). For example, when earthworms are tested in a sandy soil at low pH, the sensitivity for the contaminant may be increased due to the extra stress from the acidic soil conditions. It may be wise in such cases to remove the "extra stress" data before making an SSD for the Risk Assessment.

14.10 Other Issues in SSD Modeling and Interpretation

14.10.1 Comparison of Hazard Indices and PAF

After setting a quality standard, sites may be evaluated against that standard, yielding insight in the presence or absence of exceedances in different soil samples (a dichotomous outcome). In soil quality assessment, one can moreover derive a quantitative Hazard Index, as ratio of the actual soil concentration and the standard. When the index exceeds the value of unity, then this indicates a *regulatory* problem, signifying potential risks, but the resulting number (Hazard Index >1) does not provide information beyond the number of times the criterion is exceeded, usually interpreted as degree of seriousness of contamination. Especially if one compares sites with different contaminants on the basis of quantitative indices, one should be aware of the underpinning of the indices. Are they based on an SSD, or on e.g. a lowest NOEC divided by an uncertainty factor of 10, 100 or 1000? Usually, this is information that is hidden when quality standards are used by practitioners. This matters a lot, since the latter implies that the indices for different compounds are of different meaning. Moreover, since the maximum impact is 100% of species affected, index values are not a good approach to quantify expected impact magnitudes. The reasons for this are illustrated in Fig. 14.10, and further discussed by Klepper et al. (1998) and Solomon and Takacs (2002).

As an improvement, PAF-values may be used instead of Index values. The advantages of using PAF instead of an Index in Conventional environmental assessments

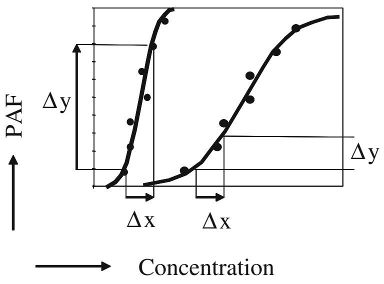


Fig. 14.10 Different contaminants will exhibit different shapes and positions for their SSDs, as shown schematically. An increase of concentrations in the environment (ΔX) implies different increases in PAF (ΔY)

are that the former has an improved ecological interpretation. While an Index (or a summed Index over compounds) has no upper boundary, the upper limit of the PAF (and net PAF for mixtures) is 100% of species possibly affected. Furthermore, PAF correctly addresses the increase of risk (ΔY for a given ΔX) when slopes of the concentration-effect relationship differ between contaminants, and doesn't suffer from the problems of different (hidden) uncertainty factors. Finally, it has been subject to validation studies (see Section 14.5).

When the regulatory context asks for follow-up actions in the case of exceedance of soil quality standards, practitioners must simply execute the prescribed activities, such as determining the urgency of remediation, given the local soil use. When an exceedance does not lead to prescribed action, expected impacts can better be quantified and ranked by estimating the local PAF.

14.10.2 Dealing with Natural Background Concentrations

It has sometimes been observed that protective soil quality standards are lower than the natural background concentration, as a consequence of neglecting Exposure Assessment details. In this case, implementing the risk-based soil quality standard would imply that such sites are considered "unacceptably contaminated", in the sense of "there is a man-made cause, we have a problem". Care should be taken to look at the variability of natural background concentrations in an area, region or country, before a risk-based protective soil quality standard is set and used for defining the presence of man-made environmental risks and a need for subsequent Risk Management. Soil concentrations may result from anthropogenic or natural sources (e.g., surface metal ores), or from both. In this case, SSD-based Risk Assessments must be combined with an appropriate Exposure Assessment to address this phenomenon. In the case of naturally elevated concentrations such as heavy metals in soils from near-surface ores, it may be wise to manage risks by preventing sensitive and valued organisms to be exposed, while not considering remediation. In the specific case of surface metal ores, remediation would be called "mining"! And remediation would lead to the local extinction of protected "metal flora". When, however, elevated metal concentrations result from human actions (or even an environmental "criminal act") Risk Management approaches, and eventually remediation, could be an option.

The techniques that can be of help to discriminate natural background concentrations from man-made enrichments is a subject beyond the scope of this chapter. See Chapter 1 by Swartjes, (this book) or Vijver et al. (2008) or Spijker et al. (2008) for more detailed information. However, handling knowledge on natural background concentrations in combination with SSDs can be done with the so-called added-risk approach, which enables the determination of the added risk due to anthropogenic enrichment, given a local natural background concentration. For technical details see Struijs et al. (1997).

14.10.3 The Influence of Soil Type and Soil Properties

The differences in soil types and soil properties among sites influence sorption of contaminants to the soil matrix, and – hence – bioavailability (see Chapter 16 by Hodson et al., this book). When acknowledging this phenomenon, regulatory agencies may need a system for soil-type dependent correction of soil quality standards, so that the standards calculated for a soil with a certain total contaminant concentration are different when soil composition (and thus sorption strengths) vary.

A system that addresses soil type differences has been in use for a long time in the Netherlands (VROM 1987), using soil-type and compound specific formulae. In practice, a formal national soil quality standard value of (e.g.) 10 mg/kg_{dw} for contaminant X can, by applying these formulae, be recalculated into soil-type specific values like 8.3 mg/kg_{dw} for soils with a high sand content, and 12.9 mg/kg_{dw} for soils with a high organic matter content. The latter (soil-type corrected) standards are then used in practical assessments like a Hazard Index evaluation. In practice, this means that the first soil would be considered "slightly polluted" when containing 8.4 mg/kg_{dw}, and the latter "non-polluted" when containing 12.8 mg/kg_{dw}.

The corrections known so far have been described as formulae that are based on the observation that different soils contain different background levels of metals, which likely is a consequence of different geochemical sorption properties. However, the formulae have been implicitly used as if the corrections made by these formulae also imply quantitatively similar differences in biological availability (see Chapter 16 by Hodson et al., this book). The latter would imply differences in actual exposures and impacts of the same total concentration. Though the quantitative corrections work out in this way (lower sorption implies a lower local soil quality standard), the system was not built as a bioavailability correction system. The system of formulae currently used in the Netherlands, developed in the 1980 s, is rather out-of-date. A modernization could result from recent scientific investigations (Spijker et al. 2008).

14.10.4 When Soil Concentrations are Very High

It is relevant to consider the level of contamination at a contaminated site, and the context of the assessment, before one decides to use NOECs, EC50s, or other measures of effect as SSD input data.

For Criterion Assessments, it may be helpful to set remediation triggers in such a way that they relate to clear types of impact, such as the loss of species. In such cases, the remediation trigger may be better estimated from an SSD based on EC50s or LC50s (Lethal Concentrations) than from NOECs. However, the choice of endpoint is a matter of policy and will vary among governments. In the Netherlands, a remediation-related criterion is based on the HC50 of the SSD based on NOECs.

The preference for an SSD-EC50 or SSD-LC50 for the appraisal of risks in highly contaminated soils relates to the fact that the SSD for a contaminant based on EC50s is plotted to the right of the SSD for that same contaminant that is based on NOECs (see Fig. 14.11 for a comparison of the two SSDs for an aquatic test system, and see also Fig. 14.6 on the comparison of HC5-NOECs with HC5-EC50s). This logical phenomenon implies that the SSDs differ with regards to the upper-tail concentration at which the estimated PAF-values do not increase any more at increasing exposure concentrations. In the figure, the fraction of species exposed beyond its NOEC is approx. 95% at an ambient concentration of 1 mg/L, so that output from such "chronic" SSDs would become meaningless *for ranking expected impacts* at concentrations of 1 mg/L or higher. This is not (yet) so for the SSD-EC50, where the expected acute impact levels of exposures of 1 and 10 mg/L still differ (with estimated Potentially Affected Fractions of approx. 60 and 90%). Moreover, an estimated fraction of species beyond their EC50 is expected to imply serious consequences for those species, whereas an estimated fraction of species exposed beyond their NOEC has no such intuitive interpretation.

The example and intuitive reasoning suggest that one should consider the soil concentrations of the site of concern when selecting appropriate test parameters to

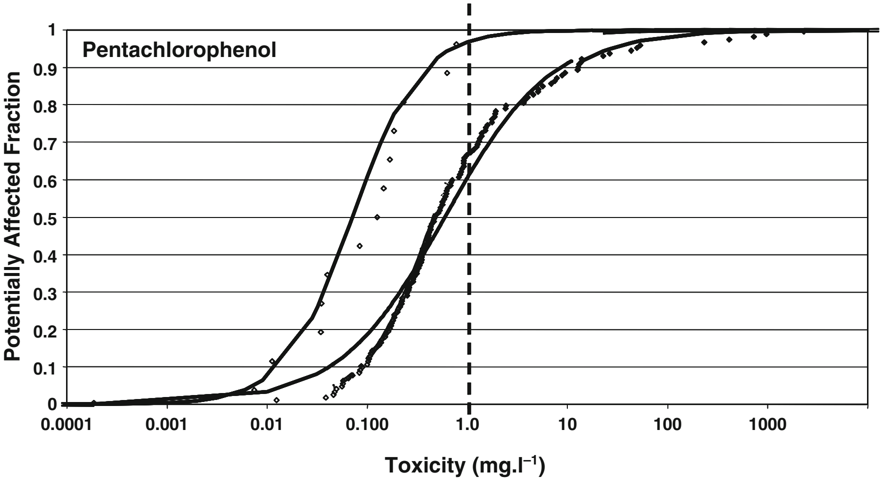


Fig. 14.11 Different ecotoxicity data for pentachlorophenol in aquatic test systems illustrate that the position of the NOECs (*left*, 13 species) and EC50s (*right*, 136 species) values used as input data for SSDs differ in position, as do the two resulting SSDs (SSD-NOEC and SSD-EC50) fitted to those data. The SSD-EC50 is to the right of the SSD-NOEC

construct the SSD. One should construct an SSD beyond the common one based on NOECs when all output of an assessment would be obtained in the upper tail of that SSD.

14.10.5 When Soil Concentrations in an Area Vary

For contaminated sites, it is common to see a huge variance in contaminant concentrations among soil samples. So, there are species sensitivity differences, but also differences in exposure levels within a specific contaminated site. Hence, the output of an SSD in terms of risk levels (PAF) in such cases would deliver a *spatial distribution of Potentially Affected Fractions*. At some spots in the site (hot spots), risks are extreme, while at other spots the risks are low. Similarly, a suite of data on toxic risks may be collected in time. For example the risk reduction due to declining concentrations as a consequence of breakdown of compounds or of leaching may lead to a *temporal distribution of PAFs.*

For risk managers, it may be worthwhile to know the variability in toxic pressures within a site or area, to help focus attention on hot spots (for remediation). A temporal distribution in PAF for an Outcome Assessment after taking Risk Management actions may demonstrate that risks were indeed reduced as a consequence of those actions.

Statistically, the need to combine the variability of exposures in space or time with that of sensitivities (species) has lead to the derivation of *Joint Probability Curves* (JPC). JPCs result from combining an exposure distribution with an SSD. The results look like the graph in Fig. 14.12, where a hypothetical site is sampled at three moments in time, at multiple spots within a site: before and after remediation of the most contaminated spots, and after autonomous breakdown of organic toxicants over time. The graphs represent a Conventional Risk Assessment of the initial situation (“high”), and two Outcome Assessments (results of the remediation “moderate” and waiting “low”). Note that “an ideal remediation” would have resulted in a sharp-angled JPC-curve, with the angle in the lower right. Such a curve reads as: “there are only very few spots left with a toxic pressure of more than 5% of the species”.

14.10.6 When There is a Mixture of Contaminants

When considering contaminated sites, it is obvious that the local biota are nearly always exposed to mixtures of contaminants. Hence, a Risk Assessment should address the issue of potential cumulative impacts. How to do that has been a problem for a long time.

An often-applied approach in practice is simply not to model mixture responses, and treat each compound separately. This is, in fact, an approach of “no addition”. It is implicitly assumed that *only the most toxic contaminant* is causing the impacts.

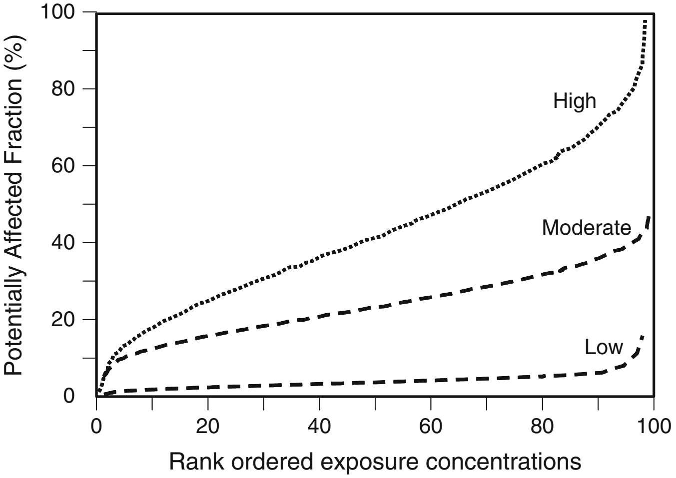


Fig. 14.12 Imaginary example of cumulative profile plots (a type of Joint Probability Curve, JPC). The three lines show the results for three sampling events: at the moment of discovery a highly contaminated site ("high") with space-varying contamination (and thus space-varying risk levels), after remediation of the hot spots within the site (resulting in "moderate" contamination and risks), and after waiting for autonomous breakdown of organic contaminants by soil microorganisms ("low": exposure concentrations and risks have dropped further due to breakdown of contaminants by microorganisms). In the "high" case, 50% of the soil samples within the site would cause effects on more than 45% of the species (high risk profile), while in the "low" case, only a few samples would affect more than 10% of the species. Graphs produced by ETX

However, this model has appeared to be wrong in a multitude of single-species toxicity tests of mixtures. An analysis of those tests, using alternative mixture models, showed that mixture impacts are most often larger than those of the most toxic compound; moreover, mixture impacts could be well predicted based on mode-of-action considerations and associated models (Altenburger et al. 2000; Backhaus et al. 2004; Hermens and Leeuwangh 1982; Hermens et al. 1984; Silva et al. 2002). Due to the observations in single-species tests of mixtures, there is no reasonable motive to expect that multi-species responses (which are the cumulation of single-species responses, though modified by species interactions) should follow the "no addition" model. Also for species assemblages, the net effect of mixtures expectedly will resemble an aggregation of the single-compounds effects rather than an effect level of the most toxic compound.

This assemblage-level impact expectation is difficult to validate. It is technically very difficult to execute empirical studies such that one can quantify mixture effects in multi-species mixture studies and to analyze those quantified responses on the basis of predicted mixtures responses as generated from (simultaneously run) single-compound effect studies and appropriate mixture models. However, the study of Pedersen and Petersen (1996) is supportive of the rule of thumb that mixture effects at the level of species assemblages can be evaluated using a mixture modeling approach.

The simplest approach in Ecological Risk Assessment of mixtures is summation of Hazard Indices over compounds. Like the use of indices for single compounds, this method has serious drawbacks. The outcomes of index summations may be meaningless, e.g., especially when different compounds have quality standards derived with different uncertainty factors.

A pragmatic approach to address mixtures in Conventional Risk assessments using SSDs was proposed by De Zwart and Posthuma (2006). These authors developed a modeling tool to handle ecological mixture risks for species assemblages, based on toxic pressure assessment. The approach originates from the models for addressing mixture impacts in basic pharmacology and toxicology, by applying Concentration and Response additivity modeling in sequence. The method is therefore a *mixed-model approach* and consists of a series of steps (Fig. 14.13). Formulae and software to run the approach are presented in the aforementioned publication.

First, all contaminants of the actual mixture (measured or predicted concentrations) are assigned to groups with similar Toxic Modes of Action, such as cholinesterase-inhibiting insecticides, photosynthesis inhibitors, and narcotics. This may not be an easy task, since one contaminant can have multiple modes of action. For example, an insecticide can be harmful to plants too. Various publications are, however, available to help assigning Toxic Modes of Action in practice, like Escher and Hermans (2002) and Verhaar et al. (1992). Many contaminants can be simply assigned to a group exerting a general narcotic Toxic Mode of Action. Despite the technical problem of assigning contaminants to groups, the outcomes of the cumulative assessment *are likely to be quantitatively robust* (see below).

Second, the net actual risk of each of the contaminants separately is derived using the SSDs for those contaminants as in Fig. 14.4 (X→Y use), with site measured (or predicted) concentrations as input (X). This results in an estimate of the PAF for each of the contaminants at the site. This step is usually preceded by an assessment

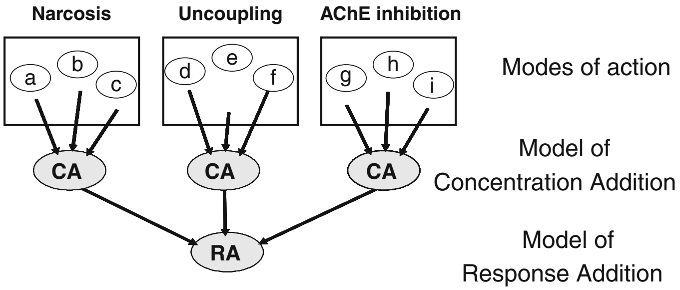


Fig. 14.13 The mixed-model approach to derive the multi-substance Potentially Affected Fraction (msPAF) at sites contaminated with mixtures. For a mixture of nine contaminants (**a**–**i**), assigned to three Toxic Modes of Action (abbreviated as: narcosis, uncoupling and AChE inhibition), the first analysis step is determining the msPAF within each contaminant subgroup using the model of CA, and then to aggregate these msPAF-CA values to an overall msPAF for the whole mixture using RA modeling. CA = Concentration Addition, RA = Response Addition

of the actual bioavailable fraction. Methods for that are described in Hodson et al. (Chapter 16 of this book), and in De Zwart et al. (2008b).

The third step is to aggregate the estimated PAF-values of each compound within groups of contaminants with similar Toxic Modes of Action. This is done by a *Concentration Additivity model*. In this modeling approach, each contaminant within a group is considered a "dilution" of the others, so that concentrations of all contaminants can be expressed as toxicity-normalized concentrations. This yields *multi-substance PAF-values* (*msPAFs*) for each mode of action represented in the contaminated soil.

The fourth step is to aggregate the within-group msPAF-values and the remaining PAF-values for contaminants not assigned to a group, to obtain an overall msPAF. This is done by a *Response Additivity model*.

As a result, a Conventional Risk Assessment for a mixture assessment using SSD-modeling yields *a single estimate of the net, Potentially Affected Fraction of species for a site*. This can be presented with or without confidence intervals. Such a single value is very important for the ranking of sites and also for the diagnostic analysis of Biomonitoring data (see Fig. 14.5). Note that this procedure, in contrast to the derivation of standards, never imposes an uncertainty factor on a PAF-estimate.

Many theoretical objections can be generated to this way of quantifying mixture risks. As for SSDs themselves, there is no real mechanistic underpinning of the approach. However, mathematical analyses of the models that are used have revealed that:

> the numerical values of mixture risks are grossly similar for the mixed-model approach, in comparison with the Concentration- and/or Response-additivity models, as long as the slopes of the concentration-response models are "moderate" (see Drescher and Bödeker 1995).

This mathematical rule of thumb suggests a certain robustness of msPAF as an estimate of expected impacts, rather insensitive to modeling choices. In line with the robust findings for single-species mixture tests, we object to assessing single contaminants one by one in retrospective Risk Assessments of contaminated sites, since the numerically robust output of the mixture models (CA, RA, or the mixed-model approach) seem (1) conceptually more justifiable than a "no addition" approach, and (2) seem better linked to field effects than contaminant-by-contaminant assessments (see Fig. 14.5).

14.10.7 When the Environmental Problem is Refined: Tiers for SSDs

The concept of tiering has been introduced in Swartjes (Chapter 1 of this book). Within the context of tiered assessment, the SSD approach can be considered a lower-tier approach. A common approach that is simpler than SSDs is to use a NOEC, and divide this by a fixed uncertainty factor (e.g. a NOEC divided by 10 or 100) to set a "safe" soil quality standard in the context of soil protection. In comparison to the concept of uncertainty factors, SSDs are more refined.

Within the SSD-context itself, however, two tiers can be considered:

1. a *basic use* of SSDs (one well-fitted model for each contaminant); and
2. a *refined use* of SSDs.

In the *refined use*, the risk assessor considers the risks of contaminants for various subgroups of organisms. That is: the input data set is split in two or more subgroups before fitting two or more SSDs. For example, in the case of soil contamination with an insecticide, one can make an SSD for arthropods and one for "other species". There are three reasons for following a refined approach: (1) the problem definition requires that the assessor follow the refined approach, (2) additional data or insights trigger the refined approach (e.g., a specific Mode of Action is identified) and (3) statistical findings like misfit of models to the data show the basic use to be inadequate.

The *refined approach* is illustrated in Fig. 14.14. A *basic* assessment for the photosynthesis inhibitor atrazine would suggest a moderate fraction of *all* species to be impacted (approximately 75%), with a considerable misfit of the model to the data. A refined assessment shows an appropriate fit of the three submodels to the data sets, and suggests that the major fraction of primary producers (estimate: 100%) would suffer from the same exposure. This leads to the subsequent *ecological* implication, that the whole system would be at risk, due to the selective risk for the primary producers followed by indirect effects on species depending on those primary producers. For contaminants with non-specific Modes of Action (like narcotic action), the refined assessment matters much less, as shown in the examples on benzene.

SSDs are quite simple as compared to higher tier methods such as food chain modeling (in which one considers transfer of a contaminant through a food chain from plant eater to top predator) or population or food web modeling (see e.g., Forbes et al. 2001, 2008). But, in contrast to the latter methods, SSD-approaches are versatile, there are many data available to use SSDs in practice, they are easy to use, they quickly yield insights that may be helpful directly in Risk Management or for focusing a next assessment step (see below), and they have an intuitive interpretation (fraction of species affected).

14.11 Weight-of-Evidence and Tiered Use of SSD Output

SSD output can be used as a *single line of evidence* or in a *weight-of-evidence* setting (multiple lines of evidence). This is treated in detail in Rutgers and Jensen (Chapter 15 of this book), where the toxic pressure output of SSD-modeling is used with toxicity tests of contaminated soils and field observations, to assess actual risks of high contaminations.

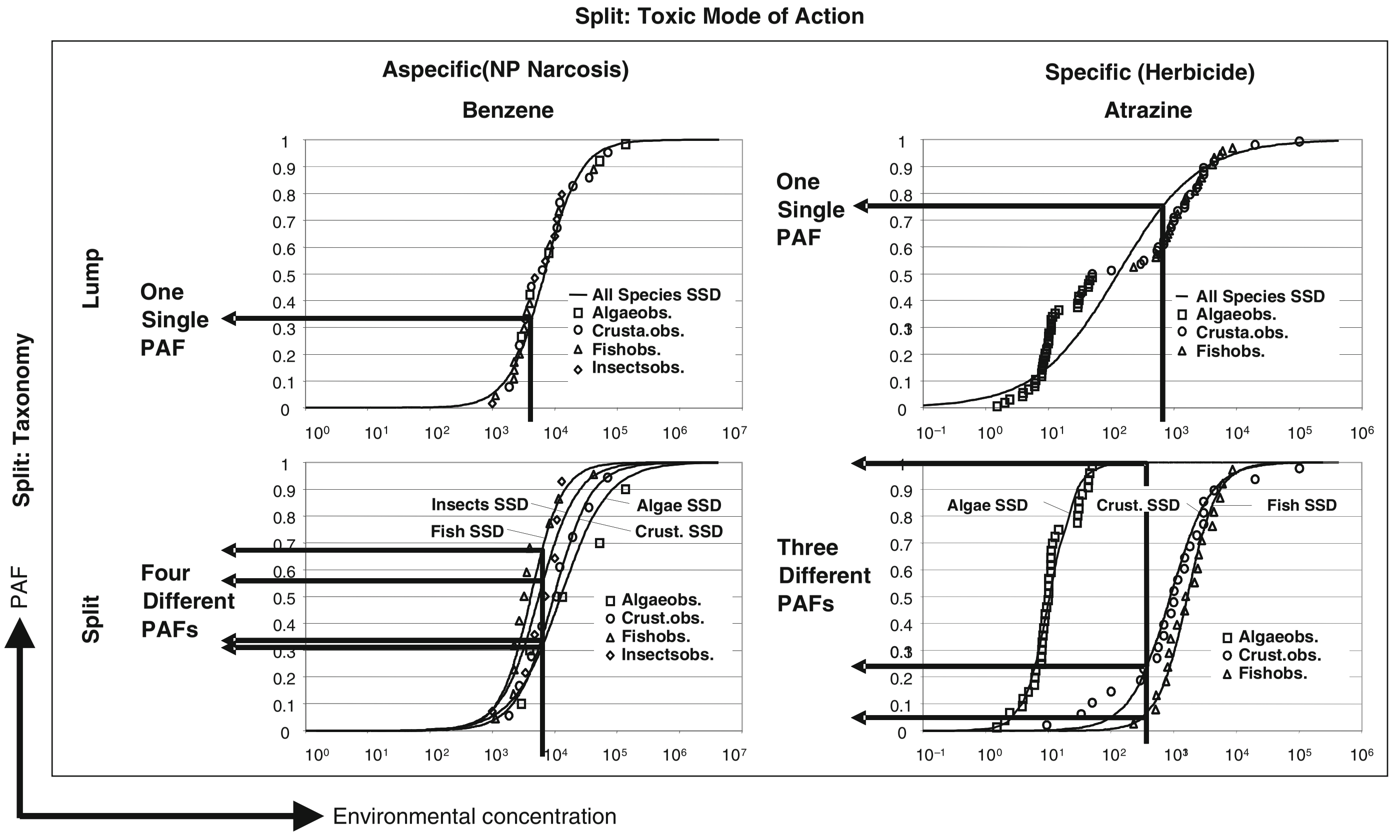


Fig. 14.14 A refined assessment with SSDs: by splitting the input data set into subsets of data for target- and non-target organisms, the assessor takes into account that contaminants may have a specific Toxic Mode of Action. The result is a refined view on potential impacts on species groups. (NP=non-polar, crust.= crustaceans, obs.= observations). Reproduced from Posthuma et al. (2002a)

This weight-of evidence approach is commonly used in tiered assessments. Higher tiered assessments are applied when the uncertainty in the SSD-output is so large, or the costs of possible interventions are so high, that one wants more refined assessments. For example, when a set of 100 sites is ranked using 100 calculated values of msPAF, one may have sufficient budget to remediate only 10 sites. Hence, one may select a top-20 of sites using SSDs only, and then apply the more costly weight-of-evidence scheme of Chapter 15 by Rutgers and Jensen, this book to give a detailed and experimentally based ranking of only those sites. This is more cost-effective than weight-of-evidence testing of all 100 sites.

14.12 Key Strengths and Limitations of SSDs

SSD-modeling has strengths and weaknesses, especially in comparison to other methods.

First, SSDs are easy to use, requiring only simple statistical software and appropriate ecotoxicity data. Such software and data are available. Second, SSDs combine ecotoxicity data for a contaminant in a way that is easily interpreted, namely in terms of a potentially affected fraction of species. Third, these estimated fractions can be aggregated to a single estimated toxic pressure for cases of contamination with mixtures; this is of key importance to solve various problems in practical Risk Assessment. Fourth, such toxic pressures appear to be related to field impacts, in the sense that increased predicted toxic pressures relate to increased species loss and impacts (though much more evidence is needed), despite the influence of other stressors that affect biotic communities. Fifth, SSD-based outcomes appear to be robust, in the sense that they can be used for relative ranking an array of compounds, and arrays of contaminated sites. Sixth, SSD-modeling can be used for very different environmental problems; a key strength is *versatility*: the concept can be used for two major purposes, derivation of formal soil quality standards and site-specific Ecological Risk Assessment. The latter can be done on the basis of measured or expected concentrations. This permits the comparison of alternative Risk Management scenarios to the existing conditions. Seventh, the major strength may be that SSDs can be of help to make Risk Assessment and Risk Management processes cost effective. Rather than 100 experimental tests, or 100 higher-tier population or food web models for 100 cases, SSDs can help to quickly select the top 20 of risky sites. This will cause a substantial cost reduction when the higher-tier assessments are done for only the selected sites, and the Risk Management activities are applied to sites with clearly significant risks.

The *limitations* of the SSD-approach relate mainly to the issue that the model is statistical and contains no ecological interactions. Critics are right when they say that the *method* has eco-deficiencies, but they would be wrong if they concluded from that fact that SSD-*output* has no practical use in Ecological Risk Assessment and Risk Management (see use examples below). SSD-output appears to result in consistent patterns of relative rankings, both amongst contaminants and amongst contaminated sites. This means that the weaknesses apparently do not invalidate the *relative* interpretations of SSD-output (the rankings).

Regarding the limitations, we acknowledge that various issues in the use of SSDs are conceptually “not solvable”. For example, whatever the refinement of an approach, it is conceptually impossible to protect all of the unknown species in the soil by any approach or model, *because* they are unknown (except for a no-emission approach). No single approach can be imagined in which a sensible solution can be provided to “the unknown”. Hence, all conceivable models and approaches are weak when extrapolating from conventional laboratory data to untested contaminated sites. This is not a weakness of SSDs, it is a universal weakness.

It is our opinion that SSDs are currently among the best risk-ranking methods *that can be practically applied at low cost to many practical Ecological Risk Assessment problems*. They primarily help to explore the risk problem in quantitative terms. In particular, they can be applied in scenario analyses, such as estimating the relative risks from predicted exposures under different Risk Management scenario’s (e.g., Fig. 14.12). When one wants to be sure about impacts at sites already contaminated, the experimental and Biomonitoring approaches as described by Rutgers and Jensen (Chapter 15 of this book) will be of use, to circumvent the key assumptions of SSDs. However, such approaches are not useful *before* an event has happened (one cannot test a field effect before a contamination event occurs, and thus cannot prevent damage of a possible event in this way), as when one needs to decide on deposition of slightly contaminated sediment on land (see Section 14.14.3).

The strengths and weaknesses of SSDs may be summarized as:

- SSDs are versatile; they have proven to be useful tools for practical soil appraisal and protection policies and for the evaluation of existing contaminated sites;
- however, the SSD-approach is not a panacea for all ecological soil contamination problems, as appropriate, higher-tier models and approaches should be applied;
- the strengths of SSDs are fundamentally related to the idea that one can rank relative risks, amongst contaminants and amongst contaminated sites (different PAFs);
- when needed, the output of SSDs can be used as decision criteria (like a soil quality standard);
- the output of SSD-based explorations can be used to design a next step (tier) in Ecological Risk Assessment in a cost-effective way (e.g., focus testing on the most likely impacted hot spots, or the contaminant with likely the highest impact); and
- for many practical problems, the versatility and the low costs of exploratory SSD modeling may outweigh scientifically sound but more difficult and expensive alternatives.

14.13 Practices of SSD Use

14.13.1 Practical Approaches in this Chapter

The examples in this chapter are all based on fitting a log-Normal distribution model to the data, except when stated otherwise. The software used in the examples is *ETX*, which is freely distributed upon request (Van Vlaardingen et al. 2004). Data for the examples were obtained from *the RIVM e-toxBase*, a storage and retrieval system for ecotoxicity data, currently containing more than 188,000 entries, representing more than 5000 contaminants and more than 2000 species tested in water, sediment or soil.

14.13.2 Criterion Risk Assessments, the Oldest Use of SSDs

A review of the presence of Risk Assessment elements in soil policies in Europe showed that at present only three EU countries (Germany, Finland and the Netherlands) have an approved guidance policy on soil Ecological Risk Assessment (Carlon et al. 2007). These three countries also include ecologically-based critical soil concentrations in their soil quality standards. The review also showed that several other countries (Sweden, the Flanders region in Belgium and Denmark) developed ecologically-based critical soil concentrations, but these values were not yet established as formal soil quality standards. SSDs may be used for deriving standards, but the minimum required number of species varies between jurisdictions. In Italy, the UK, the Walloon region of Belgium and the Czech Republic, ecological soil quality standards are in development. There is, according to this overview, political interest in ecological soil protection in the EU Member States.

The endpoints (types of test data) considered for use in deriving soil quality standards vary amongst nations, as shown in Table 14.1 (from Carlon et al. (2007)).

Soil Quality Standards are easy to use in practical soil protection frameworks. They may be used to assess whether the risks from emissions would surpass the safe standard, and to assign sites as being (seriously) contaminated. In many cases, however, soil contains contaminants for which a standard is not available. When a

Table 14.1 The ecological receptors for which (test) data are considered in the derivation of soil quality standards vary amongst some studied nations (data from Carlon et al. (2007)); standards are not necessarily derived with SSDs

	Micro. process	Soil fauna	Plants	Above-soil ecosystem	Aquatic ecosystem
Austria			X		
Belgium (Walloon)	X	X	X	X	X
Belgium (Flanders)	X	X	X		
Czech Rep.	X		X		
Germany	X	X	X	X	
Spain	X	X	X	X	X
Finland	X	X	X	X	
Netherlands	X	X	X	X	
Sweden	X	X	X	X	X
United Kingdom	X	X	X	X	

standard is lacking, soil quality assessments can be based on provisional soil quality standards, that is: soil quality standards that are not formally accepted, but that reflect the best available knowledge on the contaminant. Using provisional quality standards, one could also calculate a provisional Hazard Index. If sufficient ecotoxicity data are available, however, one can use those data also to determine the SSD, and subsequently quantify the toxic pressure associated to the concentration of the contaminant.

For certain contaminants, namely those which often occur together, it is technically possible to derive group soil quality standards, when the composition of such mixtures of contaminants is known and more or less constant. This has been done, for example, by Traas (2003) for Dutch standards for the sum of ten different PAHs in soil. Such standards simplify the assessment and regulation of mixtures.

After a contaminant is allowed to be produced and used based on an evaluation of expected exposures and sensitivity data, an Outcome Assessment could be based on an analysis of soil quality monitoring data. When the Risk Assessment models appropriately predict real-world phenomena, no exceedances of the soil quality standards should be found when the product is used. If exceedances are nonetheless found, appropriate management actions can be derived and taken, and a further Outcome Assessment would show whether those actions were effective.

14.13.3 The Dilemma of Conservative Quality Standards

Criterion Risk Assessments are used to set and implement soil quality standards. In setting such standards, authorities should be well aware of a potential dilemma that may arise when protective soil quality standards are applied: they may trigger far more regulatory actions than anticipated or they may elicit large public concerns.

Posthuma et al. (2008) have described this dilemma. On the one hand, it was explained that the use of soil quality standards in environmental management has been highly successful. Many contaminants have been banned or their emissions have been reduced, thereby reducing toxic pressure on soil communities. The success of this approach is most easily seen in aquatic systems where the ecological status has improved considerably over the last decades, in large part due to standards-based bans and regulations.

But on the other hand, the lower the protective soil quality standards or remediation trigger, the larger the areas of soils considered “contaminated”, and the higher the number of “candidate remediation cases” in an area, respectively. The general public might consider cases that are not clean as being “dangerous”, and this would in turn require substantial communication on the risks of slight soil contamination. In short, the more conservative the bias in standard setting, the bigger the public concerns and the Risk Management problem may be.

So, due to this dilemma, low soil quality standards are good for soil protection, but bad for optimizing environmental management and for the public perception of existing risks. As a general solution, *tiered approaches in soil appraisal* should be designed and implemented, to improve assessment accuracy and to better inform the public when generic standards are exceeded (Posthuma et al. 2008). The simple alternative, a general "relaxation of the soil quality standards to avoid the dilemma" is not an option, due to the success of soil quality standards in preventive soil policies.

14.13.4 From Criterion Risk Assessment to Conventional Risk Assessment

Due to the application of soil quality standards in practice, a three-part discrimination of soil contamination problems has resulted in many countries:

- "clean" (or non-polluted) soil, in which the low-range soil quality standard which was designed in the context of soil protection is not exceeded; this signals no unacceptable risks, and generally no public concerns;
- "contaminated" soils, in which the protective criterion is exceeded for one or more compounds;
- "highly contaminated" soils, in which the soil standards trigger serious concerns with respect to local risks.

The further content of this chapter concerns the latter two classes. In these situations, SSDs are used to quantify local toxic pressures, and to use the outcomes from such assessments for Risk Management.

14.13.5 Conventional Risk Assessments with SSDs: A Versatile Approach

Conventional Risk Assessments are done when there is a current or an expected source of a contaminant (or a contaminated site), and an estimate of the type and magnitude of effects (with or without confidence intervals) is required.

For Conventional Risk Assessments we need to combine SSD-modeling together with exposure modeling, including bioavailability assessment (see Chapter 16 by Hodson et al., this book) and mixture modeling (Section 14.10.6), to take into account local conditions. Combining these methods results in *a single estimate of the net toxic pressure of site contamination* (or predicted impacts of future contamination levels, given e.g. current emissions), quantified as msPAF. Such net risk output has three uses:

1. the msPAF-value can be compared to a policy-adopted maximum-acceptable *fraction of species* or *fraction of functions* that might be affected to determine whether the mixture elicits toxic pressures lower or higher than that effect-related limit;
2. it can be used to rank sites within a list of suspected risky sites on the basis of overall toxic pressures; in this case, software can be developed to handle large sets of soil concentration data (examples are provided below); or

3. it can be used in the tripartite approach introduced in Rutgers and Jensen (Chapter 15 of this book); in which all three techniques focus on net mixture effects (msPAF, next to toxicity tests of site soils and field observations).

Detailed examples of the use of toxic pressure estimates in Conventional Risk Assessments and management of contaminated sites are specified in the next Section. Here, we first list some examples (partly on aquatic assessments), *to elicit further thought and creativity*. They are meant to illustrate how soil policy formulation and soil Risk Management can profit from the SSD concept.

Toxic pressure estimation has been used as follows:

1. Focus on *Which Contaminant*? An exploration was made of the net mixture impacts of upstream production of High Production Volume Chemicals or pesticides in downstream mixing zones (the sea), see Harbers et al. (2006) and Henning-de Jong et al. (2008); the net risks in the sea appeared to be relatively low, and they appeared mainly attributable to only a few contaminants.
2. Focus on *Which Contaminant*? An exploration was made of the spatio-temporal net potential impacts of pesticide use (> 100 compounds) in a region with water bodies adjoining the croplands; this resulted in identification of the seven most hazardous contaminants on a landscape scale, which appeared linked to only two crops of key interest, potatoes and flower bulbs (De Zwart 2005).
3. Focus on *Where?* A GIS-mapping of soil, water and sediment quality in terms of toxic pressure, was used to explore the locations of sites or areas of concern, and to discriminate them from areas where (mixture) risks are unlikely to occur. This helps to focus management actions where they are needed most (cost-effectiveness), as promoted by Verdonck et al. (2003).
4. Focus on *Which Stressors, including mixtures*? An exploration was made to diagnose the relative role of environmental mixtures as compared to the impacts of other stressors in species loss in natural systems based on Biomonitoring data (De Zwart et al. 2006; Mulder et al. 2005). This is a very relevant use of SSDs, since current policies not only have looked at exceedances of soil quality standards for selected contaminants, but also have introduced a *Good Ecological Status* as a “holistic” policy target as in the EU-Water Framework Directive (European Commission 2000).
5. Focus on *Industrial Product safety by Life Cycle Assessment.* SSDs may be used to derive toxic effects metrics when assessing environmental impacts in the Life Cycle Assessment of, e.g., industrial products (Huijbregts et al. 2002) or pesticides (Van Zelm et al. 2009), to support the design of products with low environmental impacts during their *production – use – waste* life cycle.
6. Focus on *Decision Support in repetitive, complex situations*. Scenario-based Risk Assessments can support Risk Management after a natural disaster when the disaster may cause or has caused the release of large amounts of chemicals (Van Dijk et al. 2009). In this case, SSDs based on EC50-data are used to estimate the perimeter that would show 50% of species loss after a major chemical accident, which is a relevant measure for addressing likely impacts on human food sources (e.g., fish in lakes) and biodiversity.

This list of examples illustrates that SSD-based conventional Risk Assessments are used in a wide variety of practical contexts. The list should allow the readers of this book *to envisage the versatile use of SSD-based results* in the *environmental assessment and management problems that they may encounter, so that they can develop their own specific approach.* We suggest that the use of SSDs should be considered for novel problems in contaminated site management, by analogy to the suite of examples presented above. To support this more practically, a few soil related examples are worked out below.

14.14 Examples of Conventional Risk Assessment of Soil Contamination with SSDs

Case studies on the evaluation of soil contamination with SSDs were selected mainly from the Netherlands. This choice is justified, because no other nation has so much experience with SSD-based assessments of soil toxicity.

To understand the cases and their reason for being, a limited introduction to the Dutch national policy framework is provided. This introduction suggests that the early Criterion Risk Assessments yielded a policy framework that itself triggered the need for novel Conventional site assessment methods, which could be based in part on SSDs. Thereafter, the case studies are described, to illustrate the way in which SSDs are considered helpful in practical policies. As a comparison, contaminated site assessment and management for the US are described. This single comparison between two regulatory contexts shows that a different policy context can result in highly different approaches, while still fundamentally based on the same Risk Assessment framework.

14.14.1 Policy Framework Backgrounds – The Netherlands

Soil Quality Standards have been derived and applied in the Netherlands since the 1980s (Van Straalen and Denneman 1989; Van de Meent et al. 1990). Two standards were chosen: the Target Value (a level indicating "good quality"), and the Intervention Value (a level triggering a refined assessments to determine the remediation urgency (Swartjes 1999)).

The implementation of these standards appeared to result in the three types of soils, as expected (clean, moderate and highly contaminated soils). What was not expected was that national inventories thereafter resulted in hundreds of thousands of "cases of serious soil contamination" (Kernteam Landsdekkend Beeld 2004), and that the moderate level of soil contamination (Target Value < local concentration < Intervention Value) triggered large public concerns due to the vast areas where this occurs. Moreover, due to the absence of formally accepted tools to appraise the soil quality in these cases, many societal activities that involve soil use or soil transfer became the subject of fierce debates. Both problems finally resulted in a large innovation in regulations on soil protection and management (while keeping in place the regulations preventing compound loading in soils). Since 2008 the new soil policies have focused on (at least) “fitness for use”, and this required a refinement of conventional Risk Assessments both for moderately contaminated sites and for candidate remediation sites. These refinements were implemented as a Risk Toolbox for soil quality assessment (see example below), and by sub-selecting those sites where there are current risks for man, and where there are plans for area development. The new policies for managing moderately contaminated soils are described by various authors (Boekhold 2008; VROM 2007). Sub-selections of seriously contaminated sites, based on current land use and on various criteria and considerations other than contamination levels alone, recently resulted in lower numbers of sites of highest relevance (Versluijs et al. 2007), and a policy target to handle those cases within a few years.

The toolbox for refined soil quality Risk Assessment is provided to local competent authorities (www.risicotoolboxbodem.nl/english). This toolbox contains Risk Assessment formulae (such as SSDs), various default values (such as the two parameters for the log-normal SSD of a contaminant), and various regulatory standards, related to different land uses. The more sensitive a land use, the lower the exposures allowed. The toolbox is used to explore the question whether area-specific soil quality standards can be derived, tailored to the actual (exposure) situation. By using the toolbox, local soil authorities can assign the soil to a soil quality class, and assess the kind of risk imposed by current diffuse soil contamination and the magnitude of that risk. Specifically, local authorities can calculate contaminant-specific toxic pressures (PAF) and mixture toxic pressures (msPAF), to obtain a refined spatial impression on the actual risks of soil contamination for ecosystems (next to those for humans and for agricultural product quality) and to identify the most potent soil contaminants. Additionally, the toolbox allows local authorities to take into account area specific parameters, such as the specific type of land use, bioavailability and soil characteristics.

SSDs are used in all processes whenever possible: in deriving soil quality standards (HC5 and HC50), in looking at actual ecological risks of mixtures (msPAF of sites), and in determining the urgency of remediation.

14.14.2 GIS Mapping of Soil Quality

14.14.2.1 Problem Setting

Various authorities, such as municipalities and district water boards, have responsibilities in the management of slightly contaminated soils. They need to know where contamination occurs, what the contamination level means in terms of risk and Risk Management (highly contaminated sediment must be remediated), what the sources are, and what they could or should do about it. The target of such authorities may be set in terms of “a reduction of the frequency of exceedances of the soil quality standards for contaminant X within a time period of Y years”, or a “reduction of the net mixture risks". Outcome Assessments of Risk Management activities could consist of monitoring the predicted reduction of soil contamination and risks over time. An ideal case of successful management would lead to a reduction of the frequency of sites where a soil quality standard is exceeded and thus a reduction in the toxic pressure at a range of sites.

14.14.2.2 Approach

For regional authorities, soil contamination problems can be presented using GIS (Geographical Information System) maps. This is helpful to find cases of high contamination, to explore the presence and location of diffuse or point sources of contamination, and to take preventive or remedial action when relevant diffuse or point sources are found. Apart from presenting concentration maps or maps showing exceedances of soil quality standards, the distribution of local risks can be presented in terms of a map of toxic pressures (PAF per contaminant, or msPAF for mixtures). To further summarize a GIS-map in a summary risk diagram, one can use the Cumulative Profile Plot concept of Fig. 14.12. A successful risk reduction strategy would show up as a flattening of the Cumulative Profile Plot (towards the "low" curve in the figure) over time.

14.14.2.3 Conventional Risk Assessment Results

GIS-maps can present concentration levels, exceedances of quality standards, and various ways to express toxic pressures. If the total soil concentration can for example be split into a "natural background concentration" and a "human-induced enrichment part", one can map the added risk caused by human activities such as sources of diffuse emissions. The latter are most relevant in helping to deriving realistic and (cost) effective emission reduction techniques.

An example of results of this kind is provided in Fig. 14.15 for lead. Lead has been used in gasoline for decades, and has been used in building materials and other man-made products. As a consequence, a diffuse contamination pattern exists on top of many hot spots.

The figure presents maps from samples in rural and nature areas (no hotspots) with (1) total topsoil concentrations of lead, (2) total subsoil concentrations of lead (from samples not influenced by human activities, and (3) the enrichments of the topsoil per site (resulting from using a so-called baseline modeling (Spijker et al. 2008). The enrichment data relate closely to the 0.43 M HNO_3-extractable fraction (Spijker et al. Subm.); this procedure extracts the chemically reactive fraction of lead from the soil matrix (including the lead in the pore water). This reactive fraction has in turn been used to quantify the toxic pressure of lead associated to the enrichments only; this toxic pressure is thus based on the upper estimate of the potential availability of lead. This toxic pressure map (4), based on NOEC-data, suggests spatial variability in the chronic toxic pressure of lead on local soil ecosystems across the Netherlands as a consequence of human-induced enrichment. A vast majority of the samples (>>90%) is characterized by a chronic added toxic pressure below 10%.

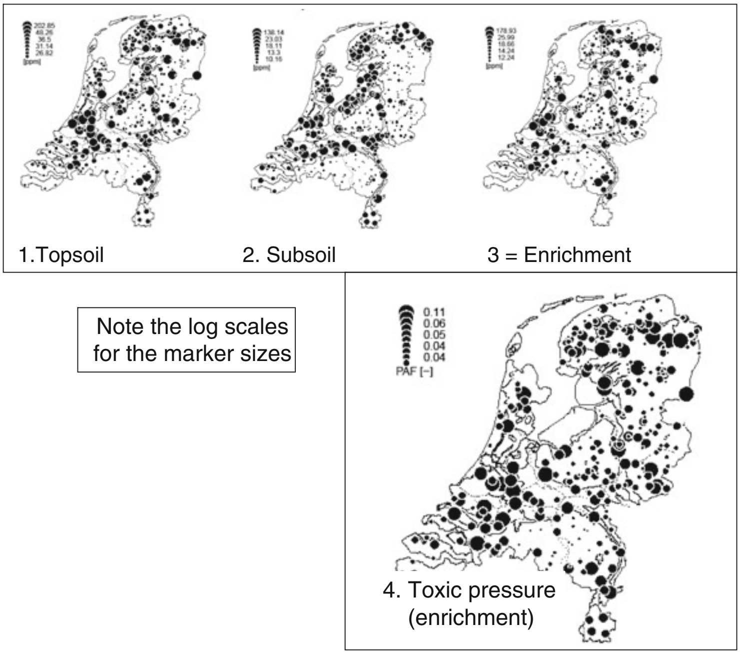


Fig. 14.15 GIS-mapping of (1) total concentrations of lead in the upper layers of Dutch soils, (2) of natural background concentrations in sub-soils, and (3) of the human-induced enrichment of lead (all mg/kg_{dw}). The toxic pressure of lead enrichments of the topsoils is shown in map 4. Fluv. clay = fluviatile clay; Mar. clay = marine clay

That is, to the best of our knowledge, the diffuse contamination with lead (on a national scale) is an unlikely cause of species loss (due to lead alone!), since there is only a low frequency of cases where the NOEC of soil organisms is exceeded. On a more local scale, modeled enrichments may be substantial (map 3), but (partly due to the sigmoid shape of the SSD) this need not be accompanied by equal increments of toxic pressure (compare subfigures 1, 3 and 4). Weaker extraction techniques (like 0.01 M $CaCl_2$) could further be used to mimic the actual mobile fraction (e.g., the metal concentration currently present in the pore water), which could result in lower estimates of toxic pressure than those obtained with nitric acid extraction. Note that the national inventory of diffuse contamination does not overlap with the inventory of sites with serious contamination (see Section 14.14.5).

14.14.2.4 Management Assessment

A Management assessment based on results like those obtained for Pb could support a decision to reduce emissions or implement regular monitoring of lead (or other contaminants) on a regional or national scale. Emission reduction measures are taken or can be taken for various metals from various emission sources (like agriculture). For example, in the Netherlands, gasoline no longer contains lead. In the future, a reduction in metal levels in the topsoil may be expected due to leaching, for lead an extremely slow process. Associated to that, the toxic pressure in the topsoil may decrease slowly. Such an effect may be counteracted by continuing diffuse inputs.

14.14.2.5 Outcome Assessment

An Outcome Assessment related to emission reduction measures has not yet been made, since the processes that change lead concentrations in the topsoil are very slow, and there is no data yet on subsequent sampling periods to allow for time-trend analyses.

14.14.3 Handling Slightly Contaminated Sediments

14.14.3.1 Problem Setting

The Netherlands is a country rife with water. Land parcels are lined by thousands of kilometers of actively managed waterways. Waterways tend to accumulate sediment, and sediments tend to accumulate contaminants. Sediments pose problems to water quantity management, so that sediments need to be removed regularly.

Given the occurrence of (slightly) contaminated sediments, policies have been formulated in the 1980s and 1990s to prevent soil contamination which would result from the old practice of depositing (slightly) contaminated sediments on adjacent lands, to act as fertilizer. At the time of implementing those policies, it was expected that preventive chemical and emission permit policies would be sufficiently effective in reducing the accumulation of contaminants in newly formed sediments in the future. In short, it was expected that all newly formed sediments could be spread on adjacent land again in the year 2000.

However, the policy ban that was formulated on spreading slightly contaminated sediment from ditches in the rural landscapes on land (based on sediment quality standards) resulted in increased sediment loads in waterways: water managers could not get rid of the slightly contaminated material that was continuously formed (AKWA 2001). This seriously threatened water *quantity* management (“keeping the country dry”), and possibly also affected water quality. In addition, spreading on land costs (say) 3 €/m^3, while removal from the system and off-site treatment or storage costs up to 100 €/m^3. The question is thus: how to balance water quantity and quality management with soil quality protection and cost effective Risk Management?

14.14.3.2 Approach

As a solution, a database containing the volumes of backlog sediments with their contamination levels was created, and the risks of spreading these sediments on land was evaluated in scenario studies. The database was used to explore not only which fraction of sediments could be (cost) effectively spread on land using the "old" standards-based assessment (based on five classes of sediment contamination), but also to devise a "new" assessment system, based on a systems-analysis in combination with toxic pressure assessment. This means that, in addition to contaminant toxicity (as in the chemical regulations), the behavior of the contaminants in the soil-groundwater system was also studied. That is: contaminants that are deposited on land might be subject to aerobic breakdown and leaching (as disappearance terms), and this type of changes can be taken into account in land-deposition Risk Assessments.

A complete conceptual model was made, to identify the relevant pathways of contaminant behavior in the system, with or without deposition of sediments on land (Fig. 14.16, top). Implementing the model with relevant parameter values yielded systems-based Predicted Environmental Concentrations, given regular deposition scenarios, with repeated removal of sediments, for example every five years. These PECs were determined for all typical deposition scenarios, and the results (in terms of predicted accumulation of contaminants in soils, and in terms of toxic pressures) were compared to various regulatory soil protection targets.

14.14.3.3 Conventional Risk Assessment Results

Figure 14.16 (bottom) presents the variability in the change of net mixture risk for all the sediment loads in the inventory. The graph represents the change in toxic pressures in the terrestrial soil. That is, the contamination level of the terrestrial soil prior to sediment deposition was taken into account according to data inventories like those shown in Fig. 14.15, and the delta-toxic pressure was calculated, showing the increase or decrease in net risks after sediment deposition. It presents the toxic pressure change after a long time of regular sediment deposition on land (Y-axis) versus the assignment of contamination classes in the older classification system. In the figure, the implications of using the "old" classification system versus the "new" systems-based approach are explored, using the expected toxic pressure increase per site as a benchmark for expected risk levels in terrestrial soils.

Apparently, "clean" sediments of the former class-0 do not induce increases in toxic pressure in terrestrial soils; a slight decrease of net risks was predicted. Further, the higher the "old" class, the higher on average the increase in net toxic pressure of mixtures in terrestrial soils. However, in each of the old-system Classes, the variation in expected toxic pressures between volumes of sediment in the inventory list was shown to be large. For example, some moderately contaminated "old" Class 3-sediments appeared less hazardous than various "old" Class-2 sediments. Upon inspection, these Class-3 sediments were contaminated with only a few organic contaminants with a high breakdown potential (causing the low expected toxic pressure) and at concentrations near the lower Class-3 limit, while the "worst" Class-2 sediments could contain a multitude of compounds that exhibit slow breakdown and/or at concentrations near the high Class-2 limit.

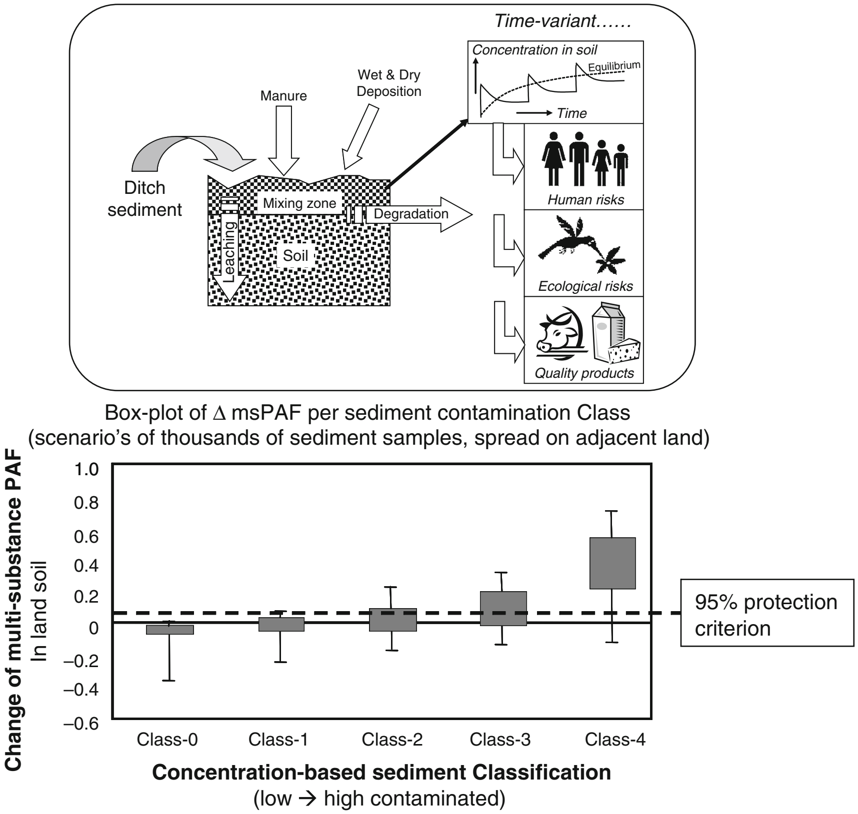


Fig. 14.16 (*Top*) The conceptual model used in a systems-based approach to address the sediment disposal problem in the Netherlands. All arrow terms were quantified, and the expected accumulation of compounds in soils and the associated net toxic pressures in soil were predicted for every sediment volume in the national sediment loads inventory database. (*Bottom*) The predicted soil concentrations of contaminants, after repeated sediment deposition on land, caused a change (delta) in the chronic toxic pressure of the mixtures in each of the volumes of a large inventory list (1356 cases, based on SSD-NOECs). Sometimes, spreading of sediments (particularly class-0-sediments) on terrestrial soil (using the old classification system) caused decreased rather than increased toxic pressure in terrestrial soils, when the terrestrial soil was more contaminated than the sediments. Based on net mixture risk predictions, a new sediment management policy was developed

14.14.3.4 Management Assessment

A suite of sediment management options were explored, including studies involving the costs of different management scenarios (Posthuma et al. 2006a, b; Van Noort et al. 2006). Since the volumes of sediments per site are known, and since the costs of various handling options are known, it was possible to explore the total costs of different management options. Based on these insights, a new policy for managing contaminated sediments has been formulated and implemented (Oste et al.

2008). By adding information on sediment management costs for each of the handling options, the cost-effectiveness of the alternative classification systems could be explored.

14.14.3.5 Outcome Assessment

Since the removal of sediments was considered urgent to prevent flooding, the new policies were implemented before the system was validated. A real *Outcome Assessment* could consist of a regular monitoring of terrestrial soils, in a chronosequence, to study whether the new policies result in terrestrial soils that are less contaminated than the soil quality standards, and/or result in lower net risks than predicted. Due to the urgency of the problem given the water quantity management needs, the Management Assessment is currently being conducted. Hence, there is as yet no Outcome Assessment showing that the new approach is both sufficiently safe as well as cost-effective. However, Harmsen (2004) demonstrated that long-term deposition of slightly contaminated (Class-2) sediments on adjacent land did result in slight but significant increases of, e.g., PAH-levels in top soils. However, the net toxic pressure did not increase proportionately, due to the sigmoid association between PAH-concentration and ecological risk. It appeared that the observed concentration increases were taking place at the level of the near-horizontal lower tail of the SSDs of the different PAHs.

14.14.4 Soil Quality Classes and Local Risks to Manage Local Soils

14.14.4.1 Problem Setting

As mentioned in Section 14.14.1, a toolbox for site-specific Risk Assessment has been provided since 2008 by the national Dutch government, to assist local soil authorities in soil management. This toolbox is applied for complex situations, such as management of the reuse of excavated soils in built areas. This example shows the soil use classification in a municipality, demonstrating that different soil uses do imply different sensitivities related to those soil uses (e.g., man can be more sensitive than ecosystems), which could consequently imply the use of different soil quality standards for appraising soil quality implications. So far, all appraisals used the national standards (Target and Intervention Values), which were usually based on the most sensitive endpoints. Since 2008, three soil quality classes have been implemented in the soil appraisal system (a class with sites with soil concentrations below the background concentration, and a class “residential” and a class “industry”). Class boundaries were derived so that the local soil use is not at risk, with higher class boundaries going from “background” to “residential” to “industry”. Further, the example shows a hypothetical outcome of the use of the Risk Toolbox for such soil use areas, given certain degrees of contamination.

14.14.4.2 Approach

The first steps that need be made according to the current Dutch soil appraisal approach are to define land uses and land use zones, and to compile contamination levels for each zone. An example of the different land uses within the area of a municipality (Gouda, the Netherlands) is given in Fig. 14.17 (top).

A municipality can thus discriminate between different land uses and different soil qualities, and check whether the current or future land use is "at risk" given the contamination levels present or expected. If risks appear, due to a mismatch of soil use and soil contamination class, the responsible authorities can start a process to alter land use or to reduce risks in other ways. The Risk Toolbox supports this evaluation, since potential mixture effects are accounted for in the assignment of soil quality classes.

After the soil use zonation map is made and compared to soil contamination class boundaries (concentrations), land use and soil contamination may be consistent (i.e., actual land use is not hampered by exceedances of land use related risk limits) or not.

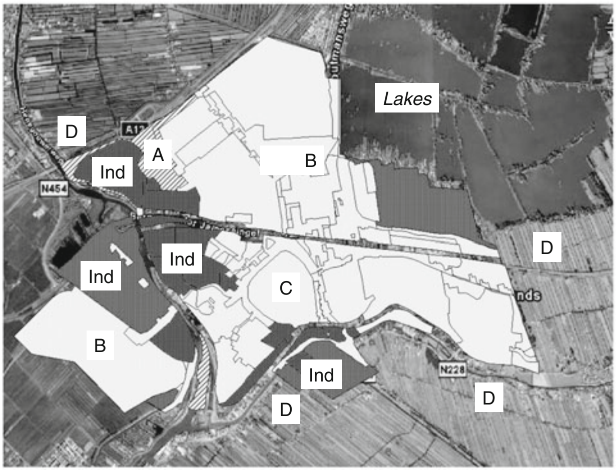
In the latter case, the exceedances of the soil-class limits does not yet specify *which* exposure scenarios would result in acceptable risk. This can then be further investigated with the toolbox and subsequent Risk Management of the existing situation.

An additional step is triggered when there is a need for soil transfer, such as during building activities. To support soil transfer, a soil transfer exploration matrix is made, looking at sources and possible sinks of soils in an area. Soils can be transferred to a site in the same quality class (e.g., "residential") or to a site with a higher permissible contamination (e.g., "industry"), so that there is no increase in risk (as could, for example, occur when transferring soil from a site in the class "Industry" to a site in the class "Residential").

14.14.4.3 Conventional Risk Assessment Results

The Risk Toolbox is of help to local soil managers if there is concern about current use of a soil zone in relation to exceedances of the soil quality standards. In that case, the Risk Toolbox provides insight into the local magnitude and type of risk, such as in Fig. 14.17 (bottom) using a Hazard Index approach.

Thereafter, the toolbox was used to quantify the local chronic toxic pressure on soil biota in a soil use with moderate ecological protection (M), in the soil use zone "Residential". The toxic pressure of the local mixture was calculated as 21% (msPAF-NOEC). This implies that the mixture exposure probably results in less protection than the 95%-protection criterion. Such a toxic pressure level may be considered by local authorities as not extremely serious for a residential area (msPAF-NOEC does not necessarily imply species loss), for example because the level is at least lower than the 50%-toxic pressure level used earlier in the remediation policies (the HC50 used as trigger implied a chosen PAF of 50% or higher for remediation, per compound).



	Human Risk Assessment					Ecological Risk Assessment			
	Hazard Indices					Hazard Indices			
Compounds	**Soil Use**	**A**	**B**	**C**	**D**	**Soil protection level**	**High**	**Middle**	**Low**
As		0.11	0.02	0.02	0.00		0.40	0.30	0.15
Cd		0.68	0.06	0.00	0.00		**1.20**	0.43	0.12
Cr		0.03	0.01	0.01	0.00		0.32	0.25	0.09
Cu		0.07	0.01	0.00	0.00		**1.24**	0.38	0.24
Pb		**5.31**	**1.39**	**1.02**	0.21		**9.02**	**1.28**	0.37
Hg		0.06	0.00	0.00	0.00		**4.62**	0.30	0.11
Ni		0.01	0.01	0.01	0.01		0.35	0.29	0.12
Zn		0.26	0.02	0.00	0.00		**7.49**	**5.21**	0.85

Fig. 14.17 (*Top*) Soil use zones for an example municipality in the Netherlands. When a soil zone is used as residential area, the soil quality for all contaminants in the soil should be safe for that type of land use, and fall within the boundaries of the soil quality class "residential". This is easily checked per compound, but more difficult when there is a need to account for mixtures; in such cases, the Risk Toolbox helps to implement the assessment steps. (*Bottom*) The Dutch Risk Toolbox refines the Risk Assessments based on generic soil quality standards (Target Value and Intervention Value). For this imaginary example, the toolbox specifies the local Hazard Indices that result from local contaminants in relation to the different land uses. Left: human health-based Risk Assessment. Right: ecologically-based Risk Assessment. Land use codes are: A = vegetable garden, B = residential, C = residential, no garden, D = nature areas. Local risks of a contamination level for humans are present in residential areas, since the toxicity criterion of unacceptable human risks was exceeded by a factor of 1.39 and 1.02 for residential areas with and without garden, but highest for the soil use "vegetable garden" (due to which exposure through food consumption is highest). The ecological Hazard Index at moderate protection for soil ecosystems (M, a protection level chosen for residential areas) is highest for zinc, then for lead, then for copper. The toxic pressure in this soil sample (msPAF) is 21%. H = high protection (soil use: nature, or agriculture), M = moderate protection (residential), L = low protection (industry)

14.14.4.4 Management Assessment and Outcome Assessment

When the standard way of assessing risks appears insufficient, e.g. in the case of large inventory lists of soils to be transferred or high levels of diffuse contamination or enhanced natural background concentrations, a municipality can choose to set *local* soil quality standards, using the Risk Toolbox. In that case, the local soil quality standards are selected so that they do not pose harm to the current or planned future use of the soil, while solving practical and urgent soil transfer problems as in the sediment management example. In this way, competent authorities are supported in local Management assessments.

For both Conventional and Management assessments (and in the future, Outcome Assessments), the Risk Toolbox has been used since mid-2008 as an instrument tailored to the Dutch policy framework. Many authorities apply the soil classification approach, whereby one determines whether soils are generally fit for particular uses. Evaluation of local toxic pressures, which is available as an option, can be used more widely.

As yet, there are no management or Outcome Assessment results. The Risk Toolbox is now being widely used by competent authorities to refine the Risk Assessments that were so far based on national, protective soil quality standards and remediation trigger values. Often, this refinement has suggested that a *current* soil use does not pose risks, despite the "warning signal" provided by the exceedance of the generic standards. The tiered system, with the Risk Toolbox being the refined tier as compared to the generic standards, appears to be of help in solving many local slight soil contamination problems.

14.14.5 GIS-Mapping of Remediation Sites and Monitoring of Remediation Policies

14.14.5.1 Problem Setting

In the Netherlands, a national inventory of cases of "serious soil contamination" resulted in a large number of sites for which (a) the Intervention Value was exceeded, and (b) a minimum-volume criterion was exceeded (Fig. 14.18, left). The Dutch government has asked for a cost-benefit analysis of various remediation policy options. The immediate regulatory problem is to select those cases where risks are highest. Another trigger to consider remediation is of course a practical trigger, like the fact that an area will be restructured for other reasons.

14.14.5.2 Approach

The analysis of costs and benefits of remediation options was explored using toxic pressure analyses and human health risks (not treated in this chapter), to assess costs and benefits of alternative remediation scenarios. The results were reported by Rutgers et al. (2006). The focus was on the contaminants that pose the highest hazards, so as to explore the need for specific remediation techniques.

14.14.5.3 Conventional Risk Assessment Results

An assessment of the contaminants posing the highest risks resulted in the graph of Fig. 14.18 (right). Apparently, PAH contamination in the inventory list relates

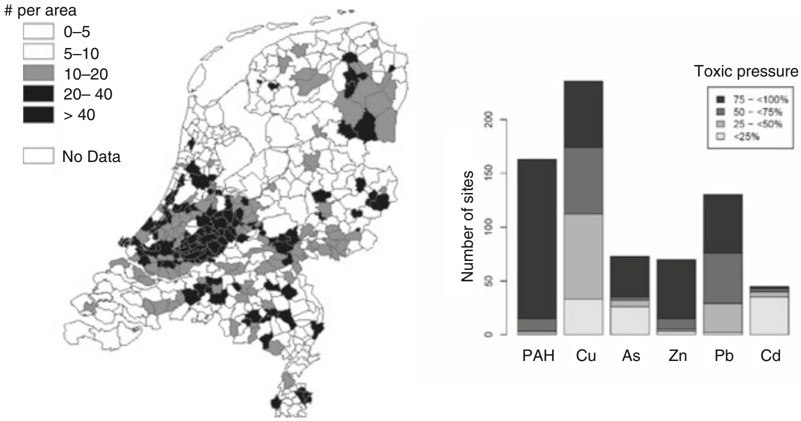


Fig. 14.18 (*Left*) Spatial distribution of numbers of cases of serious soil contamination concentration (at least one compound exceeds the Intervention Value) per area unit. This map represents more than 430,000 cases of high soil contamination, with high numbers of sites in the most populated areas. (*Right*) Numbers of seriously contaminated sites (Y-axis) with an ecological toxic pressure value in the ranges of PAF ≤ 25%, 25 < PAF ≤ 50%, 50 < PAF ≤ 75% and PAF > 75% (Rutgers et al. 2006)

to very high toxic pressures at most of the sites that contain PAHs, whereas, for example, the toxic pressures from Cd were in the lower range.

Even more insight into the most relevant contaminants can be obtained when discriminating among different types of contaminated site cases. For example, when sites that were transferred in the past to create residential areas (as foundations), the highest toxic pressure for ecosystems relates to arsenic and to a lesser extent PAHs and zinc, while for other situations zinc and to a lesser extent PAH are most important (Table 14.2).

14.14.5.4 Management Assessment

Management assessments are made yearly on the whole inventory list of cases, to focus the remediation operation more and more on the most contaminated sites (i.e., with the highest risks given the local soil use), and on the sites where there is a need

Table 14.2 Median estimated toxic pressure values (msPAF in %) for different clusters within the national inventory list of soils in the Netherlands with serious contamination (Rutgers et al. 2006)

Cluster	Cd	Cu	Pb	Zn	*As*	*PAH*
City, foundation	3	43	68	86	99	88
Dumps	30	52	96	96	53	82
Large sites	2	54	90	90	59	90
Small sites	0	43	78	78	13	88

for soil transfer and management. In practice, risks for man should be solved in the year 2015 (a policy choice), which resulted in a very large reduction of the case load for urgent action. Further, the techniques described in Hodson et al. (Chapter 16 of this book) are being used to provide information from other lines of evidence to the decision makers before, during or after a remediation is undertaken. Toxic pressure assessment is one of the three approaches being used for this purpose.

14.14.5.5 Outcome Assessment

Outcome Assessments can focus on sites and on the national inventory list. As yet, there are no published Outcome assessments focusing on toxic pressure reduction. It is however clear that the remediation resulted in concentration decreases, since that is the primary benchmark to follow the success of the remediation process.

14.14.6 A Contrasting Approach, the U.S. Superfund

Ecological Risk Assessments for contaminated sites in the United States (U.S.) are performed in the framework of the contaminated sites law known as “Superfund” (formally: CERCLA, the Comprehensive Environmental Response, Compensation and Liability Act). Superfund focuses on a relatively small number of highly contaminated sites.

In 2008, there were 1255 listed sites in the entire U.S., more than 75% of which are remediated or have final cleanup plans. The process of assessing human health and ecological risks at Superfund sites often requires several years, due to the requirements for multiple stages of consultation, planning, and site-specific sampling, analysis, and assessment.

Ecological Risk Assessments for Superfund sites are performed in two stages (Sprenger and Charters 1997; Suter et al. 2000). First, a screening assessment uses hazard quotients (HQ) which are quotients of an exposure concentration (C_e) divided by a toxicological benchmark concentration (C_b). That is:

$$HQ = C_e/C_b \tag{14.1}$$

where

C_e is a site concentration that is a conservative (i.e., high) estimate of the exposure concentration, such as the maximum observed concentration, and C_b is a conservative (i.e., low) estimate of the threshold for toxicity.

Ecological benchmark values for plants, soil invertebrates, mammals and birds exposed to soil are presented at http://www.epa.gov/oswer/riskassessment/ecorisk/ecossl.htm, including references to the origin and underpinning of the C_b values in general (in the Guidance Document) and for the numerical value of C_b for different contaminants. The C_b values for a suite of compounds are based on extensive literature searches and evaluations. They are derived separately for four groups of ecological receptors: plants, soil invertebrates, birds, and mammals. Those analyses yield ecological soil screening levels, which are defined as screening values that can be used routinely to identify those contaminants of potential concern in soils requiring further evaluation in a baseline Ecological Risk Assessment. Although these screening levels were developed specifically to be used in the Superfund Ecological Risk Assessment process, other federal, state, or private environment assessment or remediation programs can use these values to screen soil contaminants and sites in order to determine if additional ecological site study was warranted. The ecological soil screening levels were not designed to be used as remediation targets. For plants and invertebrates, the derivation procedure of soil screening levels involves collection of ecotoxicity data for plants and invertebrates and determination of the geometric mean value of the set of quality-checked and adopted test data (U.S.EPA 2003, 2005).

If mixtures are of concern, a simple concentration additive model, the Hazard Index (*HI*), is typically used, as follows:

$$HI = \Sigma(C_{ei}/C_{bi}) \ . \tag{14.2}$$

If *HQ* or *HI* is greater than 1 for any contaminant or mixture, the assessment may proceed to the second stage.

Definitive assessments use site-specific information concerning biological responses including toxicity tests of site media, modeling of food web transfer of contaminants, measurements of tissue concentrations, in-situ toxicity tests, and site surveys of biota (Luftig 1999; Suter et al. 2000). This can include an estimation of the fraction of species of different types that may be affected (i.e., PAFs, although that term is not used). Ideally, multiple types of evidence are generated for each endpoint so that the evidence can be weighed and the risks can be estimated with some confidence (comparable with the concept presented in Rutgers and Jensen (Chapter 15 of this book). If significant risks are identified, remediation targets are developed and remedial alternatives are assessed.

This two-stage assessment process brackets the SSD-based method described in the rest of this chapter. That is, the quotient method itself is considerably less sophisticated than the SSD approach, but the weighing of multiple lines of site-specific evidence is more likely to accurately characterize the ecological risks. The cost and effort of the weight of evidence approach is justified at many Superfund sites, because there are relatively few of them and many are large and have significant ecological resources. Examples include the Fox River/Green Bay system, the Rocky Mountain Arsenal, the Oak Ridge and Hanford Reservations, and the Coeur d'Alene River watershed.

14.15 Reflections and Conclusions

Reflection on and comparison of the Dutch and US approaches as two selected examples shows that the use of one type of model (the SSD) can vary depending on the context. In particular, it is surprising to see that a small country like the Netherlands has identified hundreds of thousands of potentially seriously contaminated sites using this approach, while the United States has about 1250 (seriously contaminated) Superfund sites.

The enormous difference between two jurisdictions originates from different regulatory and historical contexts. While only a few tens of cases were expected in the Netherlands in the 1980s (prior to the national inventory), the factual numbers of cases of serious contamination appeared enormous. The chosen concentration- and volume criteria to define "a serious case of soil contamination" resulted in large numbers of (small) sites, which subsequently required ranking to determine (relative) urgency for remediation due to practical limitations. The numbers of cases with serious soil contamination in the Netherlands are similar to the numbers following from regulations for surface waters under the U.S. Clean Water Act, which currently list more than forty thousand waters as biologically or chemically impaired, based on chemical and biological surveys (http://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T). In contrast, the regulation of contaminated soils under Superfund is not based on surveys of soil impairment. Rather, Superfund focuses on relatively large sites with obvious contamination. In short: legal frameworks in which SSDs or other Risk Assessment approaches are used determine the choice and implementation of scientific methods, like SSD-methods.

SSD modeling is very versatile, and can be used in very different regulatory and practical contexts. The examples given in this chapter provide an overview of possible applications, including deriving soil quality standards (classical approach), identifying the most influential contaminant in a local mixture, identifying the likely most affected locations or taxa at a contaminated site, comparing effects in Life Cycle Analysis (LCA) of industrial products and directing responses in disaster assessments. Given the relative simplicity of the model, and its ease of use given software and databases to support practical use (see below), it is left to the readers to explore whether SSD-based modeling can help in resolving their soil contamination problem.

We have presented evidence and examples for the following conclusions:

- SSDs are statistics-based summary descriptions of species sensitivity data; they are not meant to explain these differences, and they do not use ecological information; nonetheless, the output of SSD-based assessments are useful for Ecological Risk Assessment and Risk Management.
- Various useful and easy-to-use software tools and databases are available, including some on the web.
- SSDs are practically relevant for Criterion Risk Assessments to derive soil quality standards (HCp) and for Conventional Risk-, Management-, and Outcome Assessments of contaminated soils, even with mixtures of contaminants (msPAF).
- The simplest and oldest use of SSDs is to derive quality standards for a contaminant, in soils or other media, which may then be used to calculate Hazard Indices (that is, calculating the ratio of local exposure and the soil quality standard, so as to indicate a potential impact and the need for a further assessment step or a Risk Management action when the resulting index value is higher than 1).

- This oldest use is validated in the sense that estimated safe levels (most often HC5s based on SSD-NOECs) are usually associated with the absence of biodiversity and functional impacts in field or semi-field studies.
- In essence, the oldest use is a method of ranking the relative toxicity of compounds for soil organisms; and this relative ranking method appears robust.
- Furthermore, newer uses show that SSDs are a versatile, easy to use modeling tool, with which a broad array of problems can be addressed in a wide array of contexts, for many contaminants and contaminant mixtures, and for both structural (in regard to the presence of species) and functional (in regard to Ecosystem Services) responses.
- In this newer use, the SSDs rank contaminated sites in terms of estimated toxic pressures of compounds or mixtures. The higher a local toxic pressure, the higher the expected impact on species reared in such soils. There is evidence that this way of ranking is robust too, despite the problem that validation studies also show large influences of other stressors.
- The output of SSDs can be used directly, for environmental decision making or in the format of soil quality standards, or as a basis to plan further data generation or assessments.
- In addition to the scientific validation of compound toxicity and site ranking, there is "validation by use"; when a method is formally used, practice could deliver output that would challenge the acceptance of the method; regular use of SSDs has, so far, not led to credible challenges to SSDs as regulatory tools.
- SSDs can be part of a tiered approach, and in that context they are among the lower-tier methods; in such methods, they may help to focus the next step in the assessment on certain sites, taxa or compounds.
- SSDs may be used in scenario analyses to explore the (cost) effectiveness of alternative Risk Management scenarios; in this sense, various studies have demonstrated the usefulness of such explorative studies.
- SSDs may predict THAT a certain magnitude of impacts is expected, e.g. on "biodiversity" or "Ecosystem Services", but not WHAT will exactly happen (like "loss of species X").

Acknowledgement The authors are indebted to Dick de Zwart, Miranda Mesman, Piet Otte, Michiel Rutgers, Job Spijker, Frank Swartjes, Kees Versluijs and Arjen Wintersen for developing the contents of this chapter, and the editor, Hunter Anderson, and two other reviewers for their stimulating comments and worthwhile suggestions on a previous draft. Although this chapter has been technically reviewed and cleared by the U.S. EPA, it does not necessarily reflect U.S. EPA policy.

References

AKWA (2001) Base document ten-years scenario sediments – focus on slugde (in Dutch). Advies en kenniscentrum waterbodems, Utrecht, the Netherlands, 120p

Aldenberg T, Jaworska JS (1999) Bayesian statistical analysis of bimodality in species sensitivity distributions. SETAC News 19(3):19–20

Aldenberg T, Luttik R (2002) Extrapolation factors for tiny toxicity data sets from species sensitivity distributions with known standard deviation. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, pp 103–118

Altenburger R, Backhaus T, Bödeker W, Faust M, Scholze M, Grimme LH (2000) Predictability of the toxicity of multiple chemical mixtures to *Vibrio fischeri*: mixtures composed of similarly acting chemicals. Environ Toxicol Chem 19:2341–2347

Backhaus T, Faust M, Scholze M, Gramatica P, Vighi M, Grimme LH (2004) Joint algal toxicity of phenylurea herbicides is equally predicted by concentration addition and independent action. Environ Toxicol Chem 23:258–264

Beek MA, Durand AM, Hendriks AJ (2002) OMEGA: step by step: from standards to effects of contaminants (in Dutch). H_2O 17:18–20

Boekhold AE (2008) Ecological risk assessment in legislation on contaminated soil in The Netherlands. Sci Total Environ 406:518–522

Carlon C, D'Alessandro M, Swartjes F (2007) Derivation methods of soil screening values in Europe. A review and evaluation of national procedures towards harmonization. European Commission, Joint Research Centre, Ispra, Italy

Cleven RFMJ, Janus JA, Annema JA, Slooff W (1993) Integrated criteria document zinc. Bilthoven, The Netherlands: RIVM. Report nr 710401028

DEFRA (UK) (2005) Development of a web-based pesticide risk assessment module for below-ground invertebrates. Report PS(2305) DEFRA, UK

De Zwart D (2002) Observed regularities in SSDs for aquatic species. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, pp 133–154

De Zwart D (2005) Ecological effects of pesticide use in the Netherlands: modeled and observed effects in the field ditch. Integrated Environ Assess Manage 1(2):123–134

De Zwart D, Dyer SD, Posthuma L, Hawkins CP (2006) Use of predictive models to attribute potential effects of mixture toxicity and habitat alteration on the biological condition of fish assemblages. Ecol Appl 16(4):1295–1310

De Zwart D, Posthuma L (2006) Complex mixture toxicity for single and multiple species: proposed methodologies. Environ Toxicol Chem 24:2665–2672

De Zwart D, Posthuma L, Pemberton E (2008a) Attributing impacts in UK-river ecosystems to probable causes. Environ Agency. Report nr SC030189/SR5

De Zwart D, Warne A, Forbes VE, Posthuma L, Peijnenburg W, Van de Meent D (2008b) Matrix and media extrapolation. In: Solomon KR, Brock T, De Zwart D, Dyer SD, Posthuma L, Richards S, Sanderson H, Sibley P, Van den Brink PJ (eds) Extrapolation practice for ecotoxicological effect characterization of chemicals. CRC, Boca Raton, FL, pp 33–74

Drescher K, Bödeker W (1995) Assessment of the combined effects of substances – the relationship between concentration addition and independent action. Biometrics 51:716–730

Duboudin C, Ciffroy P, Magaud H (2004) Effects of data manipulations and statistical methods on species sensitivity distributions. Environ Toxicol Chem 23(2):489–499

Ellersieck M, Asfaw A, Mayer FL, Krause GF, Sun K, Lee G (2003) Acute-to-Chronic Estimation (ACE) v 2.0 with time-concentration effect models. Washington, DC. EPA600/R-03/107. U.S. Environmental Protection Agency, Office of Research Development, Washington, DC

Escher BI, Hermans JLM (2002) Modes of action in ecotoxicology: their role in body burdens, species sensitivity, QSARs and mixture effects. A critical review. Environ Sci Technol 36: 4201–4216

EUFRAM (2006) The EUFRAM framework. Introducing probabilistic methods into the ecological risk assessment of pesticides. Volume 1 of the EUFRAM project report. Version 5

European Commission (2000) Directive 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. Official J Eur Communities L 327:1–72

European Commission (2003) Proposal for a regulation of the European parliament and of the council concerning the registration, evaluation, authorisation and restriction of chemicals (reach), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) on Persistent Organic Pollutants

European Communities (2003) Technical guidance document on risk assessment in support of: commission directive 93/67/EEC on risk assessment for new notified substances, commission regulation (EC) No 1488/94 on risk assessment for existing substances, and directive 98/8/EC of the European parliament and of the council concerning the placing of biocidal products on the market. Part II

European Union (1997) Council directive 97/57/EC of September 21, 1997; Establishing annex VI to directive 91/414/EEC concerning the placing of plant protection products on the market. Official J Eur Communities L 265:87–109

Forbes VE, Calow P, Sibly RM (2001) Are current species extrapolation models a good basis for ecological risk assessment? Environ Toxicol Chem 20:442–447

Forbes VE, Calow P, Sibly RM (2008) The extrapolation problem and how population modeling can help. Environ Toxicol Chem 27(10):1987–1994

Frampton GK, Jänsch S, Scott-Fordsmand JJ, Römbke J, Van den Brink PJ (2006) Effects of pesticides on soil invertebartes in laboratory studies: a review and analysis using species sensitivity distributions. Environ Toxicol Chem 25(9):2480–2489

Harbers JV, Huijbregts MAJ, Posthuma L, van de Meent D (2006) Estimating the impact of high-production-volume chemicals on remote ecosystems by toxic pressure calculation. Environ Sci Technol 40:1573–1580

Harmsen J (2004) Landfarming of polycyclic aromatic hydrocarbons and mineral oil contaminated sediments. Wageningen University, Wageningen, 344p

Henning-De Jong I, Van Zelm R, Huijbregts MAJ, De Zwart D, Van der Linden AMA, Wintersen A, Posthuma L, Van de Meent D (2008) Ranking of agricultural pesticides in the Rhine-Meuse-Scheldt basin based on toxic pressure in marine ecosystems. Environ Toxicol Chem 27(3): 737–745

Hermens J, Canton H, Steyger N, Wegman R (1984) Joint effects of a mixture of 14 chemicals on mortality and inhibition of reproduction of *Daphnia magna*. Aquatic Toxicol 5: 315–322

Hermens J, Leeuwangh P (1982) Joint toxicity of mixtures of 8 and 24 chemicals to the Guppy (*Poecilia reticulata*). Ecotoxicol Environ Safety 6:302–310

Hopkin SP (1993) Ecological implications of 95% protection levels for metals in soil. Oikos 66:137–141

Huijbregts MAJ, Van de Meent D, Goedkoop M, Spriensma R (2002) Ecotoxicological impacts in life cycle assessment. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, pp 421–436

Jänsch S, Amorim MJ, Römbke J (2005) Identification of the ecological requirements of important terrestrial ecotoxicological test species. Environ Rev 13(2):51–83

Jänsch S, Römbke J, Schallnaß HJ, Terytze K (2007) Derivation of soil values for the path 'soil – soil organisms' for metals and selected organic compounds using species sensitivity distributions. Env Sci Pollut Res 14(5):308–318

Janssen RPT, Posthuma L, Baerselman R, Den Hollander HA, Van Veen RPM, Peijnenburg WJGM (1997) Equilibrium partitioning of heavy metals in Dutch field soils. II. Prediction of metal accumulation in earthworms. Environ Toxicol Chem 16(12):2479–2488

Joosse-van Damme ENG (1984) Ecofysiological and ecotoxicological perspectives in animal ecology. Inaugural address,Vrije Universiteit Amsterdam, 23 March, 1984

Jouquet P, Dauber J, Lagerlo J, Lavelle P, Lepage M (2006) Soil invertebrates as ecosystem engineers: intended and accidental effects on soil and feedback loops Appl Soil Ecol 32:153–164

Kernteam Landsdekkend Beeld (2004) National spatial overview. Final report on the reference measurement of the remediation work supply (in Dutch). Kernteam Landsdekkend Beeld, 110p

Klapow LA, Lewis RH (1979) Analysis of toxicity data for California marine water quality standards. J Water Pollut Control Federation 51:2054–2070

Klepper O, Bakker J, Traas TP, Van de Meent D (1998) Mapping the potentially affected fraction (PAF) of species as a basis for comparison of ecotoxicological risks between substances and regions. J Hazard Mater 61:337–344

Kooijman SALM (1987) A safety factor for LC50 values allowing for differences in sensitivity among species. Water Res 21:269–276

Luftig SD (1999) Issuance of final guidance: ecological risk assessment and risk management principles for superfund sites office of emergency and remedial response. U.S. environmental protection agency, Washington, DC, pp 1–9. OSWER Directive (9285)7–28 P

Mayer FL, Ellersieck MR, Asfaw A (2004) Interspecies correlation estimations (ICE) for acute toxicity to aquatic organisms and wildlife. I. Technical basis and model parameters. Environmental Protection Agency, Washington, DC. Report nr EPA 600/R-03/105

Mulder C, Aldenberg T, De Zwart D, Van Wijnen HJ, Breure AM (2005) Evaluating the impact of pollution on plant-Lepidoptera relationships. Environmetrics 16:357–373

OECD (Organisation for Economic Co-Operation and Development) (1992) Report of the OECD workshop on the extrapolation of laboratory aquatic toxicity data to the real environment, vol. 59. Paris, France

Oste LA, Wintersen A, Ten Kate E, Posthuma L (2008) New standards sediments. Standards for the distribution and application on water-covered sediments (in Dutch). RIZA. Report nr RIZA-report no. (2007)003, RIVM-rapport 711701064, Lelystad

Pedersen F, Petersen GI (1996) Variability of species sensitivity to complex mixtures. Water Sci Technol 33:109–119

Posthuma L, De Zwart D (2006) Predicted effects of toxicant mixtures are confirmed by changes in fish species assemblages in Ohio, USA, rivers. Environ Toxicol Chem 25(4):1094–1105

Posthuma L, De Zwart D, Wintersen A, Lijzen J, Swartjes F, Cuypers C, Van Noort P, Harmsen J, Groenenberg BJ (2006a) Sludge on soil decision making. Part 1. System approach, model en real life examples (in Dutch). National Institute for Public Health and the Environment. Report nr 711701044, Bilthoven, the Netherlands

Posthuma L, Eijsackers HJP, Koelmans AA, Vijver MG (2008) Ecological effects of diffuse mixed pollution are site-specific and require higher-tier risk assessment to improve site management decisions: a discussion paper. Sci Total Environ 406:503–517

Posthuma L, Lijzen J, Otte PF, De Zwart D, Wintersen A, Cuypers C, Van Noort P, Beek M, Harmsen J, Groenenberg BJ (2006b) Sludge on soil decision making. Part 3. Modeling of the risks following sludge distribution (in Dutch). National Institute for Public Health and the Environment. Report nr 711701046, Bilthoven, the Netherlands

Posthuma L, Traas TP, De Zwart D, Suter GW II (2002a) Conceptual and technical outlook on species sensitivity distributions. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, pp 475–510

Posthuma L, Traas TP, Suter GW II (eds) (2002b) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, 587p

Posthuma L, Van Gestel CAM, Smit CE, Bakker DJ, Vonk JW (1998) Validation of toxicity data and risk limits for soils: final report. National Institute for Public Health and the Environment. Report nr 607505004, Bilthoven, The Netherlands

Risbey JS, Kandlikar M (2007) Expressions of likelihood and confidence in the IPCC uncertainty assessment process. Climatic Change 85:19–31

Rutgers M, Spijker J, Wintersen A, Posthuma L (2006) Ecological effects of soil contamination. Cost benefit analysis of soil remediation. RIVM. Report nr 607021001, Bilthoven, The Netherlands

Silva E, Rajapakse N, Kortenkamp A (2002) Something from “nothing” – eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. Environ Sci Technol 36(8):1751–1756

Solomon KR, Brock T, De Zwart D, Dyer SD, Posthuma L, Richards S, Sanderson H, Sibley P, Van den Brink PJ (eds) (2008) Extrapolation practice for ecotoxicological effect characterization of chemicals. CRC-Press, Boca Raton, FL

Solomon KR, Takacs P (2002) Probabilistic risk assessment using species sensitivity distributions. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, pp 285–313

Spijker J, Mol G, Posthuma L (Accepted for publication) Regional ecotoxicological hazards associated with anthropogenic enrichment of heavy metals

Spijker J, Van Vlaardingen PLA, Mol G (2008) Backgroundvalues and the relation with soiltype in Dutch soils (in Dutch). RIVM – Dutch National Institute for Public Health and the Environment. Report nr RIVM report no. 7117101074, Bilthoven, The Netherlands

Sprenger MD, Charters DW (1997) Ecological risk assessment guidance for superfund: process for designing and conducting ecological risk assessment, interim final. U.S. Environmental Protection Agency, Environmental Response Team, Edison, NJ

Stephan CE, Mount DI, Hansen DJ, Gentile JH, Chapman GA, Brungs WA (1985) Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. US EPA ORD ERL. Report nr PB 85–22(7049), Duluth MN, pp 1–97

Struijs J, Meent DVD, Peijnenburg WJGM, Hoop MAGTvd, Crommentuijn T (1997) Added risk approach to derive maximum permissible concentrations for heavy metals, how to take into account the natural background levels? Ecotoxicol Environ Safety 37:112–118

Suter GW (2002) North American history of species sensitivity distributions. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology. Lewis, Boca Raton, FL, pp 11–18

Suter GW, Cormier SM (2008) What is meant by risk-based environmental quality criteria. Integrated Environ Assess Monitor 4(4):486–489

Suter GW, Efroymson RA, Sample BE, Jones DS (2000) Ecological risk assessment for contaminated sites. Lewis, Boca Raton, FL

Suter GW, Traas TP, Posthuma L (2002) Issues and practices in the derivation and use of species sensitivity distributions. In: Posthuma L, Suter GW II, Traas TP (eds) Species sensitivity distributions in ecotoxicology, 1st edn. CRC, Boca Raton, FL, pp 437–474

Swartjes FA (1999) Risk-based assessment of soil and groundwater quality in the Netherlands: standards and remediation urgency. Risk Anal 19(6):1235–1249

Traas TP (2003) Evaluatie van groep- en somnormen in het kader van Integrale Normstelling Stoffen. National Institute for Public Health and the Environment. Report nr 601501 016, Bilthoven

Tyler G (1984) The impact of heavy metal pollution on forests: a case study of Gusum, Sweden. Ambio 13:18–24

UNESCO (2005) The precautionary principle. Paris, France

U.S.EPA (1978a) Federal Register, 43: 21506–2(1518) May 18. US Environmental Protection Agency, Washington, DC

U.S.EPA (1978b) Federal Register, 43: 2(9028) July 5. US Environmental Protection Agency, Washington, DC

U.S.EPA (1995) Fed. Regist. 60: 15366–1(5425) March 23. (The methodology for deriving aquatic life Tier I criteria is on pages 15393–15399, whereas the methodology for deriving aquatic life Tier II values is on pages 15399–1(5400)) U.S. U.S. Environmental Protection Agency, Washington, DC

U.S.EPA (2002) ECOTOXicology Database system. Available at http://www.epa.gov/ecotox/. Accessed 18 August 2010

U.S.EPA (2003) Guidance for developing ecological soil screening levels. November. Office of Solid Waste and Emergency and Remedial Response. OSWER Directive 92857-55

U.S.EPA (2005) Ecological soil screening levels for cadmium. Interim final. OSWER Directive 92857-65, March 2005

Van de Meent D, Aldenberg T, Canton JH, Van Gestel CAM, Slooff W (1990) Desire for levels. Background study for the policy document “setting environmental quality standards for water and soil”. Report nr RIVM-no.670101002, Bilthoven, The Netherlands

Van Dijk S, Brand E, De Zwart D, Posthuma L, Van Middelaar J (2009) FEAT – Flash environmental assessment tool to identify acute environmental risks following disasters. The tool, the explanation and a case study. RIVM. Report nr 609000001, Bilthoven

Van Noort P, Cuypers C, Wintersen A, De Zwart D, Peijnenburg WJGM, Posthuma L, Harmsen J, Groenenberg BJ (2006) Sludge on soil decision making. Part 2. Foundation of fate and transport and estimated concentrations in terrestrial soil following the distribution of sludge on land (in Dutch). National Institute for Public Health and the Environment. Report nr 711701045, Bilthoven, the Netherlands

Van Straalen NM (1993) An ecotoxicologist in politics. Oikos 66:142–143

Van Straalen NM, Denneman CAJ (1989) Ecotoxicological evaluation of soil quality criteria. Ecotoxicol Environ Safety 18:241–251

Van Vlaardingen PLA, Traas TP,Wintersen AM, Aldenberg T (2004) ETX2.0. A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. National Institute for Public Health and the Environment. Report nr RIVM-no. 601501028, Bilthoven, The Netherlands

Van Zelm R, Huijbregts MAJ, Posthuma L, Wintersen A, Van de Meent D (2009) Pesticide ecotoxicological effect factors and their uncertainties for freshwater ecosystems. Int J LCA 14(1):43–51

Verdonck FAM, Janssen CR, Jaworska JS, Vanrolleghem PA (2003) Geo-referencing of probabilistic risk of new chemicals in rivers. Water Sci Technol 48(3):39–46

Verdonck FAM, Jaworska J, Thas O, Vanrolleghem PA (2001) Determining environmental standards using bootstrapping, Bayesian and maximum likelihood techniques: a comparative study. Analitica Chim Acta 446:429–438

Verhaar HJM, Van Leeuwen CJ, Hermens JLM (1992) Classifying environmental pollutants. 1: structure-activity relationships for prediction of aquatic inherent toxicity. Chemosphere 25:471–491

Versluijs CW, Mulder HFMM, Van Wijnen HJ, Van den Broek HH (2007) Soil remediation in political priority areas (in Dutch). RIVM. Report nr 607700001, Bilthoven, The Netherlands

Vijver MG, Spijker J, Vink JPM, Posthuma L (2008) Determining metal origins and availability in fluvial deposits by analysis of geochemical baselines and solid–solution partitioning measurements and modelling. Environ Pollut 156(3):832–839

VROM (1987) Environmental program 1988–1991 Progress report (in Dutch). Staatsuitgeverij. Den Haag

VROM (2007) Policy view of 22 November 2007, regarding the rules with regard to soil quality (Act on soil quality) (in Dutch). Staatsblad 469:1–173

Wintersen A, Posthuma L, De Zwart D (2004) The RIVM e-toxBase. A database for storage, retrieval and export of ecotoxicity data. National Institute for Public Health and the Environment, Bilthoven, The Netherlands

Chapter 15
Site-Specific Ecological Risk Assessment

Michiel Rutgers and John Jensen

Abstract In many countries, soil quality is expressed in chemical concentrations as Soil Quality Standards to address the potential ecological risks in a first tier of an Ecological Risk Assessment (ERA). In cases where application of these standards do not provide satisfactory results, additional tools are required. In this chapter the focus is on these tools, i.e. ERA taking into account the complete mixture of contaminants and the integration of data from bioassays and field ecological observations according to a weight of evidence approach. A straightforward Triad framework, combining three lines of evidence, was introduced in the Netherlands in 2007 and is presented here.

Contents

M. Rutgers (✉)
National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: michiel.rutgers@rivm.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_15,

15.1 The Soil Ecosystem and Site-Specific Risk Assessment

Site-specific Ecological Risk Assessment (ERA) is a process aiming at the support of site management decisions with respect to contamination (Suter et al. 2000). Typically, site-specific ERA focuses on a specific site. A broad spectrum of decisions can be considered, such as adaptive land management, changes in land use, and tailoring the site remediation objectives. In order to arrive at these decision points, data have to be collected, organized and analyzed to estimate the risk of contamination for ecosystems. ERA encompasses a complex procedure as many issues have to be addressed. A comprehensive ERA requires various contributions from stakeholders, authorities, managers and experts at different stages of the process before it can be fully accomplished.

Site-specific ERA ranges from rather simple or small situations to very complicated processes with many experts involved and numerous data evaluations conducted, often leading to tailored decisions. The commonality of different ERAs arises from a persuasive notion of adverse ecological effects, irrespective of the complexity or dimensions of the site. Any ERA should start with the application of generic and conservative principles for optimum protection (first tier Risk Assessment). This may be accomplished, for instance, by comparing contaminant concentrations at the site with national Soil Quality Standards, which may be adjusted for differences in soil characteristics and background concentrations (a common practice in the Netherlands, see Chapter 1 by Swartjes, this book). For the majority of sites such a generic Risk Assessment is sufficient to exclude unacceptable risks. However for a number of sites the uncertainty in this kind of generic and general assessment may be too high, e.g. when the Soil Quality Standards do not provide the right insight or the Soil Quality Standards are exceeded. This will often trigger more site-specific and less-generic actions, in higher tier Risk Assessment. In this stage, divergence between experts may occur, because different investigations/disciplines may not provide similar conclusions. Divergence between authorities and stakeholders may also reveal as a result of soil – or rather land – being treated as real estate with fixed boundaries, while contamination and ecological damage typically cross such site boundaries. Therefore, ERA should be embedded in structured frameworks allowing complex paradigms to be developed and the outcomes to be transparent, uniform and applicable for contaminated site management decisions (Barnthouse 2008; Hope 2006; Linkov et al. 2006).

For ERA in terrestrial systems, lessons can be learned from aquatic and sediment systems (Chapman and Anderson 2005; Rutgers and Den Besten 2005). Terrestrial systems, however, differ because they are generally more heterogeneous, have much slower dynamics, their food web characteristics are yet unidentified, and the contamination characterization may be less strictly assessed through the complex impact of the soil matrix on ecological effects. As a result, uncertainty is a key issue that needs specific emphasis in terrestrial ERA.

For these reasons it is essential that ERA is organized in phases, or tiers, including predictive as well as descriptive methods in order to reduce these uncertainties in a practical way. The successive tiers require increased inputs and, as a rule of thumb, more time, effort and money. The paradigm for ERA in specific cases may vary considerably, but typically includes an initial problem formulation based on a preliminary site characterization, followed by a tiered risk characterization, and it ends with a list of Risk Management options.

The question of what to consider as an aspect of the ecosystem needing consideration in an ERA is not so complex as one might think. Indeed, Egler (1977) has stated: ‘ecosystems are not more complex than you think, they are more complex than you can think’. This notion automatically provides a rationale for simplification, i.e. it is justified to address only a few aspects which should be documented, rather than deliberately trying to ‘catch it all’. Consequently, it is better to report on A risk, instead of THE risk of contamination (Rutgers et al. 2000). Secondly, aspects needing consideration may vary from very broad and general to site-specific peculiarities. In the Netherlands any ERA starts with a broad and conservative assessment via application of Environmental Quality Criteria aimed at protection of the complete ecosystem. This relates to the protection of both biodiversity and ecological functions, which is obtained through the application of so-called Species Sensitivity Distributions (SSDs; Posthuma et al. 2002; see Chapter 14 by Posthuma and Suter, this book) for species and processes (Sijm et al. 2002). In addition, rather conservative thresholds are applied for protection targets and remediation targets, i.e. 95 and 50% respectively (Swartjes 1999). Thirdly, the ecosystem approach may be broken down in distinct steps, by addressing different aspects at different levels of integration. For instance, in the case of the terrestrial environment the focus might be spread over four important aspects connected to carbon and energy transmission (Fairbrother et al. 2002):

1. primary production, i.e. focusing on organisms performing photosynthesis, e.g. green algae and plants;
2. fragmentation, i.e. focus on organisms involved in the cutting and grinding of large organic fragments and organic macromolecules e.g. earthworms and micro-arthropods;
3. decomposition and mineralization, i.e. a focus on the final breakdown and synthesis of organic components in the soil e.g. micro-organisms, protozoa and worms (earthworms and pot worms);
4. consumption and predation. i.e. a focus on (the stability of) interactions between organisms in so-called trophic webs. e.g. nematodes and micro-arthropods.

So, it is defendable to restrict ERA in the earlier tiers to, for instance, these four generalized aspects. In the latter tiers it is defendable to extend the ERA by including specific species such as protected wild life (nature) and ornamental plants (parks, gardens).

15.1.1 Appreciation of the Ecosystem at Contaminated Sites

Before any site-specific investigation is initiated, it is important, as the *first step* in an ERA, to evaluate whether there are any ecological concerns associated to this specific site (Fig. 15.1, in which the *Framework for site-specific ERA* is given).

In most countries, no detailed and systematic inventory has been made of how often ecological concerns could be associated to contaminated sites. This is for example true for Denmark. Denmark has for decades collected data and generated a comprehensive and relatively complete record of the contaminated sites within the country (Danish EPA 2008). To date this inventory has registered approximately 24,000 contaminated or potentially contaminated sites. However, a screening of how frequently valuable ecosystems, e.g. *Natura 2000 areas*, are located on contaminated sites was not initiated until recently. The investigated area covered one of the five Danish Regions. Here a total of more than 600 contaminated sites were located at – or in very close vicinity of – an important conservation nature area (terrestrial, fresh water and marine) corresponding to approximately 10% of the contaminated sites in the region. The dominating sources of contamination in these areas were tar from coating of fishing nets in the late history, shooting ranges and dump sites.

A comprehensive study in the Netherlands has shown that out of 500,000 suspicious locations, approximately 28,400 potentially contaminated sites are located within recognized nature areas or *Natura 2000 areas* (Versluijs et al. 2007). It is expected that 3,200 sites in these areas have to be remediated, comprising a total surface area of about 8,400 ha.

In the Netherlands the approach to ERA might be different from many other countries with a soil protection policy. A Risk Assessment for the terrestrial ecosystem applies for all sites with a serious soil contamination, and remediation should be seriously considered for all unpaved and uncovered soil, including those at industrial sites. In this sense, the ecosystem has the same status in the Risk Assessment as human health and the chance of dispersion and spreading of the contaminants (Swartjes 1999; Versluijs et al. 2007). This policy pays tribute to the notion that soil harbors important natural functions, which are essential for mankind. Consequently, human and ecological risk may trigger remediation at contaminated sites for all land uses, albeit the thresholds differ.

15.1.2 Stakeholder Involvement

The *second and the third step* in the site-specific ERA would then be to select relevant stakeholders and experts for the steering committee and the team of risk assessors (Fig. 15.1). The size and shape of such a Steering committee and risk assessor team depends on the type and magnitude of work anticipated for the

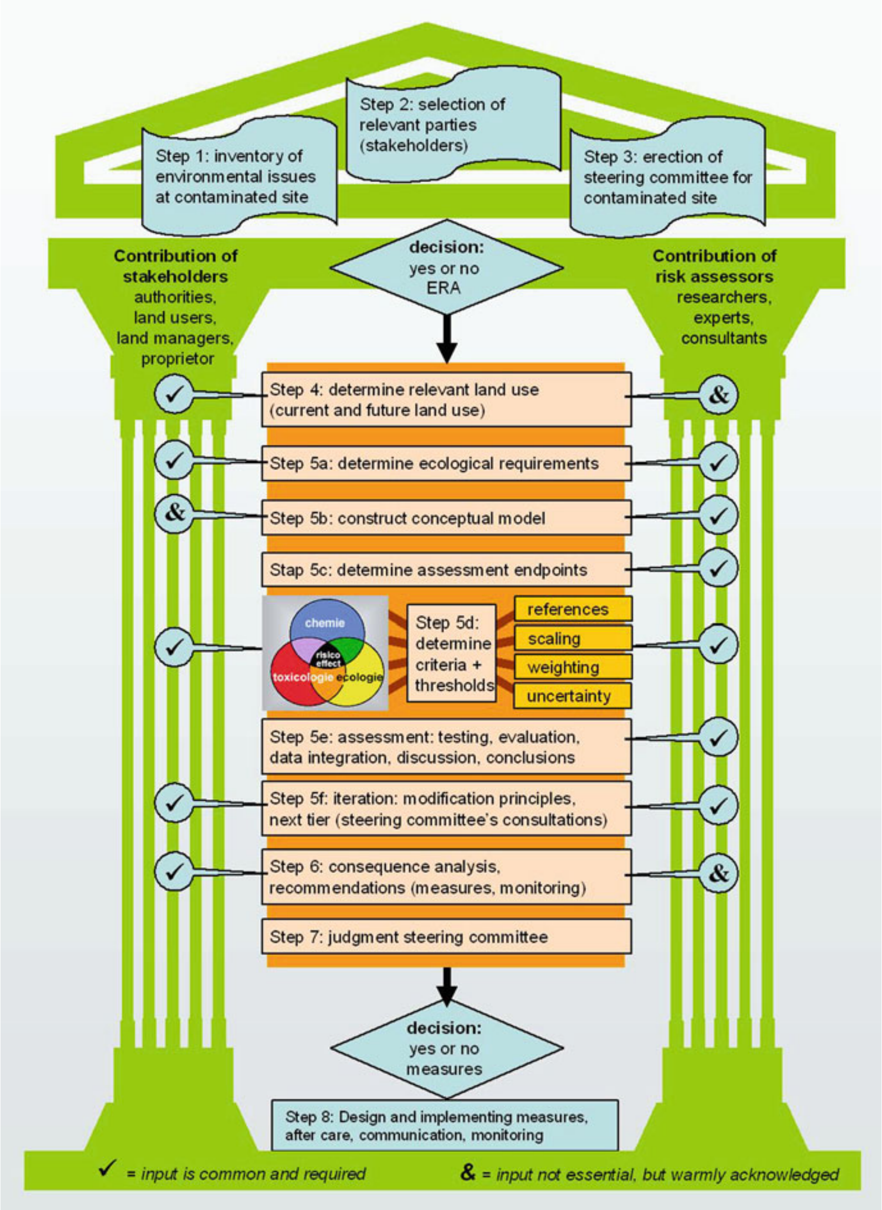


Fig. 15.1 Framework for site-specific ERA depicted in stylized portal. The pillars represent contributions from stakeholders and land users (*left*) and risk assessors and experts (*right*). After the decision to start up ERA, a steering committee should perform guidance and evaluations. The steps 4 and 5 can be used as guidance on the assessment pathway. The subsequent step 6 is included in order to facilitate the incorporation of remediation options. Step 7 comprises an independent judgment from a peer review. The framework was based on an earlier publication (Rutgers et al. 2000), slightly modified and is currently incorporated in a procedural standard of the Netherlands Normalization Institute (NEN 2010)

respective site. However, it is important to involve a wide range of stakeholders early in the process, in order to reach a mutual understanding and acceptance of the conceptual site model for the terrestrial ecosystem, including the target of protection and means of successful risk mitigation prior to initiating any actual investigations. Stakeholder involvement should therefore include contaminated-site experts from authorities, land users and land managers/owners. The team of risk assessors should include people from academia and consultancies, capable of performing ERA, covering various field of expertise.

The Steering committee should then, in alliance with the risk assessors, determine the land use as the *4th step* in the ERA. Subsequently, the actual site-specific Risk Assessment is initiated as the *5th step* of the ERA by identifying the ecological requirements related to the specific land use (5a). In the subsequent steps (5b, 5c and 5d), a listing of relevant assessment endpoints for each of the identified ecological requirements needs to be identified and agreed upon, i.e. if nutrient cycles and plant biodiversity are considered important in relation to this specific land use, a suite of tools, bioassays or monitoring end-points should address the related endpoints. Examples of relevant assessment endpoints, like sensitive crops, key species, decomposition and nutrient cycles for various land uses can be found in Jensen and Mesman (2006).

15.2 Working Hypotheses, Definition of Conceptual Models and ERA Frameworks

Contributions from and interactions by risk assessors and risk managers are essential in the definition of the conceptual model and working hypotheses. In the conceptual model a simplification of the real system is obtained in order to frame the results of the Risk Assessment. The conceptual model contains two key elements (US-EPA 1998): (i) a set of working hypotheses and (ii) a diagram representing the links between the working hypotheses. Consequently, the conceptual model sets the limits of the ERA. Terrestrial ecosystems are complex and dynamic systems. It therefore requires a well elaborated conceptual model to reduce complexity and integrate system attributes in order to develop clear solutions and management decisions. A unifying ecosystem theory is lacking, for example, making the selection of assessment end-points difficult. ERA can focus on specific endpoints, like the protection of particular species (e.g. endangered species, wild life) or the performance of Ecosystem Services of the soil system (e.g. nutrient cycling, Natural Attenuation, water retention, etc). However, ERA can in principle also cover risks derived from a more ethical perspective of environmental protection. Consequently, all biotic elements will be potentially useful to some extent.

Working hypotheses are specific assumptions about potential risk to assessment endpoints (US-EPA 1998). They are formulated on the basis of one or more information sources like contamination history and data, professional judgments and information from the ecosystem at risk. Working hypotheses are important elements in ERA because they improve the level of site specificity compared to the generic application of Soil Quality Standards. Usually, adverse effects of contaminants on ecosystem attributes are formulated in terms of the pathway, from the presence of contaminants to the potential adverse ecosystem effects, i.e. the source – pathway – receptor links. This conceptual model operates on the working hypothesis of an established pathway between source and receptor.

Conceptual models may range from very simple to rather sophisticated and complex models. A very simple and commonly used conceptual model relates to the derivation of Soil Quality Standards, where the corresponding working hypothesis is that all organisms are equally important in the ecosystem (Posthuma et al. 2002). More sophisticated conceptual models have also been used (e.g. Baird et al. 2008; Bennett et al. 2007; Faber 2006), for instance with modifications in Exposure Assessment (e.g. including bioavailability considerations) or in the end-points (e.g. field birds) and ecological processes. Regardless of the level of detail, these models will always embody a simplification of the actual ecosystem. Once conceptual models are formulated, they serve as a framework for the selection of tools and for the definition of thresholds in the assessment.

Any set of tools for ERA should be embedded in a decision-making framework, which primarily consists of phases such as initial problem and scoping phases, exposure and Hazard Assessment and Risk Characterization. Several decision-making frameworks have emerged in the literature, basically following the same outline. The US-EPA has published one of the more advanced frameworks (US-EPA 1998), including many later amendments (Barnthouse 2008; Suter et al. 2000). Also in Canada and Europe, frameworks were published (CCME 1996; Faber 2006; Jensen and Mesman 2006; Weeks and Comber 2005). In the Netherlands such a framework is recently accepted in a procedural standard (Fig. 15.1).

In this chapter we will not review and discuss various frameworks for ERA, but instead focus on a few practical issues related to the application of additional tools in a *weight of evidence* (WoE) approach.

When all these important first steps have been fully discussed, the actual Risk Assessment procedure can start. The next paragraphs will describe one of the most

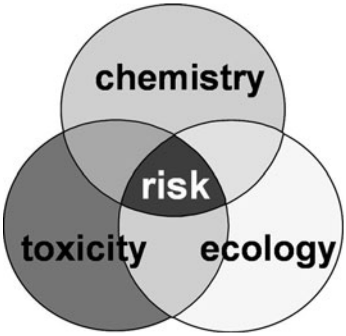


Fig. 15.2 Schematic presentation of a weight of evidence approach using the Triad. The three independent lines of evidence consist of a chemical-based assessment, a toxicity based assessment using bioassays, and an ecological assessment using data from ecological field surveys

operational and reliable methodologies, i.e. the *Triad approach* which combines three *lines of evidence* (Fig. 15.2).

15.3 Weight of Evidence and the Triad Approach

In order to deal with uncertainties in the process of ERA in a pragmatic and responsible way, it has been proposed to use a weight of evidence (WoE) approach (Chapman et al. 2002; Hull and Swanson 2006; Long and Chapman 1985; Rutgers and Den Besten 2005; Suter et al. 2000). The rationale is that multiple and independent ways to arrive at the same type of conclusions will provide a stronger 'evidence' for ecological effects, substantially improving the reliability of ERA. Unfortunately, precise definitions and application schemes of WoE in ERA are unclear (Weed 2005), but a series of papers edited by Chapman et al. (2002) addressed several issues. In this chapter we will not focus on clarifying these issues, but instead we present relevant scientific developments and practical considerations for the application of a WoE at contaminated sites (Critto et al. 2007; Dagnino et al. 2008; Faber 2006; Jensen and Mesman 2006; Weeks and Comber 2005). In addition, we illustrate these considerations by a newly adopted framework in the Netherlands (Mesman et al. 2007; Rutgers et al. 2008b).

For terrestrial ecosystems, WoE approaches are still in an exploration and developing stage (Critto et al. 2007; Jensen and Mesman 2006; Rutgers and Den Besten 2005; Semenzin et al. 2007, 2008; Suter et al. 2000). The Triad approach relates to a specific form of a WoE (Fig. 15.1; step 5d). It is based on the simultaneous deployment of three independent types of assessment tools:

- site-specific *chemical characterization* often combined with the estimation of ecotoxicological effects using literature data, e.g. by calculating a risk index;
- application of bioassays or biomarkers in order to determine de novo and ex situ *toxicity* in soil samples from the site;
- on-site *ecological observations* or other monitoring data that provide insight in the plausible effects of the contamination.

The major assumption is that WoE using a combination of tools from these three independent disciplines will lead to a more detailed and correct assessment than an approach, which is solely based on one of these, for example the total concentrations of contaminants at the site. A multidisciplinary approach will thereby help to minimize the chance on false positive (incorrectly assuming that there are effects, whilst in reality there are no effects) and false negative (incorrectly assuming that there are no effects) conclusions.

The advantage of the Triad approach can also be stated as follows: the combination of three simple instruments enables the reduction of model uncertainties, which is compatible or better than reducing model uncertainties using one sophisticated tool. Information about model uncertainties can be deduced from results of tools from different disciplines, rather than from one set of tools in one discipline. This makes the Triad approach a scientifically sound and practical instrument, during different stages of ERA.

15.4 Practical Issues for Adoption of the Triad Approach

15.4.1 Uncertainty

Uncertainty is a key element in Risk Assessment and should be properly addressed and communicated. Uncertainty can be seen as the state of imperfection within the total available amount of information with respect to the environmental problem and the requested decision to be made in time (Walker et al. 2003). Uncertainty contains both subjective and objective elements. The subjectivity originates from the judgment about the validity and appropriateness of the information. Objectivity comes from the data and facts related to the contaminated site. Uncertainty is therefore often separated in:

- Variability and error, i.e. lacking or imperfect data and data from systemic variations in space and time. An example is the variation in results by application of a specific tool: in one assay with real replicates, with pseudo replicates, from inter-laboratory variation, and through gradients in space or time.
- Incertitude, i.e. model imperfections; or in more popular terms: you do not know what you do not know. This uncertainty is demonstrated by application of different tools at one occasion (sample or site), both within a line of evidence, or between different lines of evidence.

It is important to realize that both types of uncertainty need appropriate, but inherently different approaches in the Risk Assessment. Recently, linguistic uncertainty was introduced additionally to these two types of uncertainty (Carey and Burgman 2008; Levin 2006). Linguistic or language related uncertainty between risk assessors and risk managers may arise especially in the case of ERA, because of a lack of appropriate terms and definitions, imprecise problem framing and different perspectives and views on the environment (Kellett et al. 2007). A clear and transparent communication between stakeholders before, during and after the execution of an ERA is therefore crucial: it is the only way to minimize the chance on misperceptions.

It is beyond the scope of this chapter to recapitulate all aspects of uncertainty. Instead references are made to the respective literature, (e.g. Beer 2006; Burton et al. 2002a, 2002b; Levin 2006; Nayak and Kundu 2001; Walker et al. 2003). As an idea we state that variability and error are primarily solved by increasing the amount of effort, e.g. via more samples, more replicates, and further optimizing the noise to signal ratio via improving of assessment tools. Weight of evidence approaches like the Triad seems to be preferred in order to reduce uncertainty caused by incertitude (including ignorance and indeterminacy).

15.4.2 Selection of Assessment Tools

The application of the Triad approach comprises the selection of tools which:

(1) fit in the specific tier of interest (from screening-level to highly sophisticated tools);
(2) cover all three lines of evidence;
(3) effectively address the selected end-points.

The final suite of selected tools should allow for dealing with and ultimately reducing uncertainty in ERA. A tool is defined as an instrument for quantification of a specified aspect of the ecosystem. The outcome must ultimately be expressed in a one dimensional number indicating the level of ecological effect on the uniform scale. Tools range from very simple (a screening level bioassays or a concentration plus literature toxicity data) to highly sophisticated and integrated (results from BLM modeling, maturity index of nematodes or a food web stability index). Any tool must be based on site-specific information through modeling and/or measurements and on information from literature data and ecotoxicological reasoning for interpretation of the data on the uniform effect scale (see below for an explanation in Section 15.4.3).

Elaborating on the three issues for selecting tools (see former paragraph):

- Sub 1 (tiers). With respect to the tailoring of the specific tier of interest, standardization and costs of analysis are important issues for selection of tools, especially in the lower tiers. Yet, even screening tools should be sufficiently reliable and sensitive to demonstrate effects of contaminants under field realistic conditions. Finally the tools should be relevant for the ecosystem under investigation. More sophisticated and elaborated tools are used for improving site specificity in the higher tiers of the ERA.
- Sub 2 (lines of evidence). Each tier of the Triad approach should cover three independent types of assessment tools, representing three different lines of evidence. This requires at least one tool for a chemical based assessment (*chemical characterization*), at least one bioassay (*determination of toxicity*), and at least one type of on-site ecosystem observation, which can be related to effects of contamination (*ecological observations*). When the different lines of evidence are comparable in terms of effort and matching level of insight, a balanced weighting between the lines of evidence can be applied (see below for more details on weighting of the results).
- Sub 3 (addressing selected end-points). The appropriateness of respective tools to serve as indicators for selected endpoints is the third and last issue. The tools should provide insight about compliance of end-points with respect to the potential effects of the contamination at the site. Many ecologically relevant end-points cannot be directly assessed, because of imperfect knowledge and lack of tools. Instead models or surrogate systems are used to extrapolate from the assessment tools to real world situations. Confirmation of ecological significance of the individual test systems originates from track records or literature evidence of the

respective tools in comparative cases. If not, this should be specifically addressed. This is often the case with tools used in higher tier Risk Assessment, because of insufficient scientific foundation.

As a rule of thumb, simple, common, standardized and low-cost tests should be used in the lower tiers of the ERA, whereas more laborious and sophisticated tests should be applied in the higher tiers. Guidance for selecting appropriate tools is available (e.g. Fairbrother et al. 2002; Jensen and Mesman 2006; Römbke et al. 2006a; Rutgers et al. 2008a). The highest level of protocol standardization of tests is reached in international guideline programs such as ISO and OECD. Whereas the OECD test program has focused on tests suitable for the evaluation of chemicals, the ISO guideline program has, at least recently, focused on test systems for the evaluation of the risks of contaminated soil (Römbke et al. 2006b). Additional protocol standardization comes from quality assurance systems like Good Laboratory Practice (GLP). Methods described in the scientific literature can also be used, especially in the higher tiers of ERA. In any case, it may frequently be necessary to adopt the tests to site-specific conditions. The number of laboratories able to perform a specific test on a routine or semi-routine basis is another issue when selecting the set of tools. Finally, the acceptance of the tests by the stakeholders, risk assessors and the scientific community plays a role in the selection criteria. Last but not least, it is important to minimize the bias caused by the risk assessor's background. It is, for example, a human commonality to overstress the importance of the own expertise in solving complex problems. With large multidisciplinary research teams, however, this problem is somewhat reduced.

The combination of different tools from different disciplines at different levels of standardization, robustness, sensitivity and ecological relevance without being able to fit them all in one comprehensive and accepted ecosystem theory is in fact a matter of combining 'apples and oranges'. This highlights the need for a proper *multi-criteria decision analysis* (MCDA) in ERA (Chapman et al. 2002; Linkov et al. 2006). With MCDA it is possible to combine different pieces of not a-priory related information in an unconstrained way. It opens the possibility to cope with divergent, but *best professional judgments* from separate experts in a transparent framework process.

15.4.3 Quantification and Scaling

Essentially, the results from all tools to be applied, including bioassays and ecological field surveys, should be funneled into the ERA. Key for efficient use of information is 'scaling'; i.e. the projection of results from different tools on a common and unified 'effect scale' (e.g. inhibition of growth, or loss of reproduction should be both expressed in the comparable units as an ecological effect). The primary aim is to maximize the utilization of individual results, and to use results from all tests together in transparent and integrative schemes, for example in a decision matrix. Burton et al. (2002a, b) reviewed several possibilities for disseminating final WoE findings, and concluded that tabular decision matrices are the most transparent and quantitative representations.

Ideally, ERA for aquatic, sediment and terrestrial systems should follow the same set of conventions for scaling. In practice, however, there are slight differences for the following reasons:

1. There is a wide range of standardization levels and terrestrial methods differ in sensitivity, making it difficult to define one set of homogeneous 'rules' for interpretation. Although initial thoughts for scaling of e.g. bioassays, biomarkers and community-level end-points are obtained from *best professional judgments*, still much experience is lacking. It is expected that these rules can be obtained step by step from the building up of practical experience from ERAs at contaminated sites.
2. Interpretation of test results in terms of 'effect' or 'no-effect' inevitably will result in the loss of valuable quantitative information. Except for the situation for ERA in surface water and sediment systems, the limited experience with use of the Triad approach for terrestrial systems demands for exploration and efficient use of virtual all available information in a quantitative manner.
3. In aquatic systems toxicity can be determined after a pre-concentration step, allowing the application of relatively insensitive tools and producing fewer false negative results. It is virtually impossible to concentrate soil samples putting higher demands on tools and the use of results in ERA.

For evaluation and integration of the results from the three lines of evidence in the Triad (chemistry, toxicity, ecology) a *quantitative decision matrix* is constructed. To this purpose, it is necessary to use a uniform *effect scale* for the quantification of each of the separate effect levels in the Triad approach, running from zero (no effect) up to 1 (maximal ecosystem effect). Consequently, the results from each tool (bioassay, biomarker or ecological field survey) should be projected on this effect scale, according to best available knowledge from the literature or best professional judgments (BPJ) from consulted experts. Useful and advanced examples of scaling rules and the construction of such a *quantitative decision matrix* can be found in Jensen and Mesman (2006), Dagnino et al. (2008) and Semenzin et al. (2008).

Different tools will obviously require different approaches. For instance, for a growth test the percentage of inhibition can be implicitly used as the measure for effects. For ecological field monitoring, however, the results should be scaled relatively to the ecological state of a reference site (= 0), and a (theoretical) state indicating 100% effects. Information from field monitoring is often composed from multiple variables putting specific demands on the scaling of multi-dimensional information to a one-dimensional effect value (Jensen and Mesman 2006).

Furthermore, the method of scaling should account for limitations in working range of an assessment tool with respect to the effect scale. This is sometimes denoted as the biological scale of the measurements (e.g. Gaudet et al. 1995; Wright and Welbourn 2002). The effect scale is usually defined on the level of populations of protected species, whole communities, ecosystem functions or some kind of Ecosystem Integrity (Suter et al. 2000). However, assessment tools addressing subcellular responses like biomarkers are rather sensitive and can be perfectly used as early warning signals, but have a limited range on the effect scale, i.e. a relatively low signal. On the contrary, field surveys at the population or community level are less sensitive but generally 'closer' to the assessment endpoints, making the response on the effect scale much stronger (closer to 1).

In cases with large and wide-spanning datasets, it might be feasible to apply a suite of indices in order to take advantage of all data in the best suitable way (Dagnino et al. 2008). Examples of such indices are:

- Environmental Risk Index (ERI): Quantifying the level of biological damage at the population level, comparable to the Triad with similar legs.
- Biological Vulnerability Index (BVI): Using e.g. biomarkers to assess the potential ecosystem stress and threats to biological equilibria.
- Genotoxicity Index (GTI): Used to screen for genotoxicity effects.

Whereas the first index is assigned to the ecological leg of the Triad, the two latter are assigned to the ecotoxicological leg.

These indexes were used on a site-specific case in north Italy by Dagnino et al. (2008). They showed that the Triad-based decision system (Environmental Risk Index) as well as the biomarker-based index, identified the two contaminated industrial areas as under risk. However, in contrast to the result from the Environmental Risk Index, the results from the biomarker studies, i.e. the BVI, indicated that also the chosen low-contamination site was under stress and that in some of the sampling occasions, the GTI index at this site was comparable to the contaminated industrial sites, indicating a general stress syndrome in soil organisms from that region.

These different indices allow for an elaboration on the plausible links between causes and effects. Finally, when one answer is required to aid contaminated site assessment and management decisions, these three indices should indicate adverse ecological damage on the uniform effect scale (0–1 effect scale) too.

Projection of test results on the uniform effect scale requires a certain level of experience. This expertise is *fundamental* to ERA, the importance of it can not be overlooked. Without sufficient expertise one cannot expect a responsible underpinning of the decisions from the site-specific Risk Assessment. When the issue of scaling is properly and responsibly solved, the information from separate tools from individual disciplines can be effectively used together in ERA. Fortunately, the WoE approach will help to address mismatches of specific scaling methods due to wrong assumptions (Chapman et al. 2002). Together with ecotoxicological reasoning, this information can than be used to correct the scaling method of respective tools. Accordingly, lower tiers in the Triad approach should contain tests which are, to some extent, standardized, while at higher tiers the comparative less-standardized tests should play a role in order to improve the level of site-specificity.

Once all results are quantified in the uniform effect scale, the overall response of a set of (biological) methods can be calculated. To this purpose, a weighting algorithm of different test results is required. This is described in Section 15.4.4.

15.4.4 Weighting of Effect Values

Besides the issue of scaling, the risk assessor should pay attention to the issue of weighting. Weighting applies to different tools, i.e. weighting within a line of evidence, and applies to weighting across different lines of evidence in the Triad approach. Some general principles apply to this. As a default, the three lines of evidence in the Triad should be equally weighted. Each line of evidence has its own weaknesses and strengths. However, together they form the strongest basis for ERA according to the principles of a balanced WoE approach. In specific cases, specific considerations demand for a differential weighting between the different lines of evidence in the Triad approach. The absence of adequate reference sites is typically the most problematic with ecological field surveys at strongly disturbed sites. In these cases, ecosystem changes might dominantly be caused by other factors than soil contamination. Another example of differential weighting is a difficult chemical assessment, because of complicated exposure routes and limited toxicity data. In that case it is defendable to give a lesser weight to the chemical based assessment (*chemical characterization*) than to the two effect values from the other lines of evidence (*determination of toxicity* and *ecological observations*).

Within one line of evidence attention should be given to a suite of aspects within the ecosystem. Typically, the starting point is an equal weight for all organisms and processes, applying the following popularized statement: 'All organisms are unequal, but equally important'. Another possibility is to collate data in different trophic groups like primary producers, decomposers of organic matter (fragmentation and mineralization) and consumers, and give these different trophic groups equal weights. Within any individual line of evidence of the Triad approach, differential weighting of results may be applied for three possible reasons:

1. Ecological considerations, e.g. from different land use classes, may trigger a differential weighting, which should be defined in the conceptual model. This allows extra attention to specific (functional) groups, key species, endangered species, 'charismatic' species or even specific ecological processes in the terrestrial ecosystem.
2. Differential weights can be applied in order to account for the uncertainty or variation within the end-points. Tests with a high level of uncertainty, or with a high variation in results, may be given a smaller weight in the ERA (Menzie et al. 1996).
3. Differential weights might be applied in order to correct for biases in the expected number of false positive or false negative results. For instance, the geometric mean of the inverted effect value gives extra weight to those observations with a positive response. This acknowledges the fact that many bioassays or ecological field surveys are sometimes not able to demonstrate ecological effects on the screening level, although in reality these effects are present (false negatives). This is especially a problem with tight budgets or highly dynamic systems, because the number of replicates is often too limited for demonstrating significant effects.

Den Besten et al. (1995) used differential weights in the ERA for aquatic systems following a multi-criteria decision analysis. Effects on e.g. top predators and benthos received a higher weight than parameters such as mouthpart (mentum) deformities. This information was used to rank different sites according to their possible ecological risks. For the terrestrial system, less experience is available. Semenzin et al. (2007, 2008) and Critto et al. (2007) developed tabular decision matrices to address the issue of weighting.

15.4.5 Reference Information

A crucial issue when analyzing the results of bioassays or ecological field observations in different tiers of ERA is the *reference information*. This information can be gathered from reference sites, reference samples, or literature data. Of course, analysis of reference sites and reference samples is preferred, since this optimizes the site-specificity in ERA. Due to a lack of sites and samples, literature data may partially substitute a lack of suitable references. Rutgers et al. (2008a) recently published reference data of soil system attributes for ten common combinations of land use and soil type in the Netherlands, which may be used as a source of reference information for ERA.

The issue of reference data is relevant for any line of evidence in the Triad approach, i.e. chemical characterization (i.e. background levels in that region), toxicological data from bioassays (i.e. reference soil for quantification of the no-effect level and control soil in order to verify the test performance) and ecological field surveys (i.e. the ecological status of reference sites). The perfect reference response resembles the response from the contaminated soil in all relevant aspects, besides the effects from soil contamination. When a site contains gradients in soil characteristics, also multiple references have to be gathered in order to reflect this gradient. To reach this goal, parameters that may affect test performance, like the soil's texture, pH, organic matter, humidity and available nutrients, should be verified between contaminated and reference soils. Sometimes, information is available about the influence of soil characteristics on test performance (e.g. Natal-da-Luz et al. 2008). It often is a practical problem to identify matching soil samples. This problem has to be tackled in a sensible way and hence should be considered and discussed in detail before initiating the ERA. The lack of suitable reference sites in field surveys may, however, statistically be solved by the use of multivariate techniques (e.g. Kedwards et al. 1999), which relate the species composition and abundance to gradients of contaminant concentrations in soil, taking into account possible effects of other factors ('confounders'). However, such an approach needs the analysis of large numbers of samples in order to account for all possible gradients that may shape the ecological parameters in the survey (Rutgers 2008). Many software tools are available and have increased the possibility to use powerful multivariate analysis, which use all collected data to evaluate effects at a higher level of organization. Of course, in a strict sense, causal inference of field effects from contaminants is impossible, due to imperfect reference information (Boivin et al. 2006; Everitt and Dunn 2001; Jensen and Pedersen 2006; Rutgers 2008).

15.5 Integration of Lines of Evidence and Final Results

After the results have been scaled for each test, it is possible to integrate the results of the different tests in each line of evidence. In Table 15.1 an example of collecting and presenting the data from a Triad-based ERA is given. In order to integrate

Table 15.1 Example of a table for collecting, summarizing and integrating data from a Triad-based ERA

Triad aspect	Parameter	Weight factor	Sample A	Sample B	Sample C
Chemistry	Sum TP total concentrations	1	0.00	0.76	0.92
	Sum TP porewater concentrations	1	0.00	0.62	0.75
	effect (chemistry)		**0.00**	**0.70**	**0.86**
Toxicology	Microtox	1	0.36	0.21	0.70
	Earthworm test	1	0.00	0.00	0.52
	Germination test	1	0.00	0.05	0.20
	effect (toxicity)		**0.14**	**0.09**	**0.30**
Ecology	Nematode community analysis	1	0.00	0.50	0.55
	Microbial parameters	1	0.00	0.25	0.45
	Micro-arthropod community analysis	1	0.00	0.15	0.32
	Plant community analysis	1	0.00	0.00	0.60
	Earthworms	1	0.00	0.45	0.24
	effect (ecology)		**0.00**	**0.29**	**0.45**
	Effect assessment chemistry	1	**0.00**	**0.70**	**0.86**
	Effect assessment toxicology	1	**0.14**	**0.09**	**0.51**
	Effect assessment ecology	1	**0.00**	**0.29**	**0.45**
	Integrated assessment (risk)		**0.05**	**0.42**	**0.67**
	deviation		0.14	0.55	0.38

In a first step the data are grouped per line of evidence, i.e. chemistry, bioassays and ecological field surveys. Weighting factors are set to 1 by default (first column). After calculation of one effect value per line of evidence, the data are recollected in a final set Triad data in order to judge the level of (dis)agreement between the lines of evidence (lowest tabular square). When the deviation factor ($D = 1.73 \times$ standard deviation) between the lines of evidence is low enough (see text), an integrated risk value can be used for underpinning the site management decision. In the Netherlands this lay out of the table is proposed for presenting results of a Triad as part of an ERA (Mesman et al. 2007) all data, an interdisciplinary weighting over all three lines of evidence has to be applied, which has serious disadvantages. It may be argued that as well the integration within (intra) and between (inter) the various lines of evidence in principle concerns ‘comparing apples and oranges’. However, for the moment it is the best approach available, although it is still open for improvement and adjustment.

The first integration process, i.e. within one line of evidence, aims to get a sufficient and complete set of information for estimating the risk from soil contamination. Different pieces of information are used for this evaluation. For instance, the application of Species Sensitivity Distributions (SSD) adopts the reasoning that all organisms are equally important, although they have a different sensitivity towards the contaminants at the site (Posthuma et al. 2002). Furthermore, estimates of effects based on contrasting exposure scenarios, like pore water and food exposure, may be used together to account for species-specific differences in bioavailability.

Table 15.2 Example on how to interpret the outcome of the integrated risk analyses of the Triad. It is highly recommended that stakeholders and risk assessors produce such a table before the start of the Triad process (reproduced with slight modifications from Jensen and Mesman 2006)

		Conclusion (land uses)	
Deviation (*D*)	Integrated risk (IR)	Acceptable	Not Acceptable
$D < 0.4$*	$0 < IR < 0.25$*	All land uses	–
	$0.25 < IR < 0.50$	A, R, I	N, A (with ecological and nature targets)
	$0.5 < IR < 0.75$	I, (R)	N, A, R (with ecological and green functions)
	$0.75 < IR < 1.0$	I (only with sealed soils)	N, A, R, I (with ecological and green functions)
$D > 0.4$ further investigations or (alternatively):	$0 < IR < 0.25$	A, R, I	N, A (with ecological and nature targets)
	$0.25 < IR < 0.50$	I, (R)	N, A, R (with ecological and green functions)
	$0.5 < IR < 1$	I (only with sealed soils)	N, A, R, I (with ecological and green functions)

*These numbers are arbitrarily chosen, and can be part of the negotiation process between stakeholders, authorities and risk assessors. The goal of this table is to demonstrate the common sense of choosing criteria for interpreting Triad results in the decision-making process.

D is a deviation factor indicating the level of disagreement between the lines of evidence of the Triad ($D = 1.73 \times$ standard deviation). IR is the integrated risk value from three different lines of evidence (arithmetic mean). ‘Not acceptable’ land use does not necessarily have to imply remediation or adapted soil management, but could also lead to more investigations. *N* nature, *A* agricultural sites, *R* residential sites, *I* industrial sites.

In the second and last integration step, the independent pieces of information from the three lines of evidence are compared. In this step, it is also evaluated to what extent the three lines of evidence indicate comparable levels of risks. At this point, a weight of evidence approach will pay off. Consequently, when all lines of evidence point in the same direction, it is defendable to calculate a final risk index based on the outcome of three different lines of evidence, and then compare the result with a benchmark value to be able to take a decision about the site-specific ecological risks. The benchmark value is a decided value by the stakeholders, the local administration and national government, which marks the border between acceptable and unacceptable effects (see Fig 15.1: step 5d). When the three different lines of evidence do not point in the same direction, the deviation between the three lines of evidence should be calculated and used to decide whether more research is necessary. Jensen and Mesman (2006) and Mesman et al. (2007) developed decision tables in order to arrive at these 'go/no-go decision points' to further harness a Triad approach (Table 15.2).

15.6 Embedding ERA in Formal Assessment Frameworks

15.6.1 An Example of a General Framework from the Netherlands

In the Netherlands, the Soil Protection Act was introduced in 1986. Contaminated sites are first assessed using a set of Soil Quality Standards, i.e. *Target Values* and *Intervention Values*. These values take both human health and ecological risks into account, and are applied to all kind of land uses and soil types (Rutgers and Den Besten 2005; Swartjes 1999). Recently, also so-called *Maximum Values* were introduced as remediation objectives, which are land use specific (Dirven-Van Breemen et al. 2007, 2008). The ecological basis of these Soil Quality Standards is a SSD, constructed from No-Observed Effect Concentrations (NOEC values) from the literature (Posthuma et al. 2002). At seriously contaminated sites, remediation or other soil management decisions are required if unacceptable risks cannot be refuted, based on a site-specific ecological Risk Assessment, a Human Health Risk Assessment, and the chance for dispersion of the contaminants. For these three issues, a tiered approach called the Remediation Criterion is used (VROM 2008). The first and second tiers of the ERA in the Remediation Criterion are based on a judgment of the likely ecological effects from chemical concentrations in generalized models for toxicity and mixture effects. Note that this numbering of tiers is formal and does not include the numbering of tiers in a Triad approach. In the first tier of the Remediation Criterion, the Intervention Values are used as Soil Quality Standards, besides criteria for impacted soil volume. In the second tier, ERA is performed on the basis of a calculation of the Toxic Pressure of the mixture of contaminants and a decision table addressing critical dimensions of the impacted area (Table 15.3) and presumed land use sensitivity for contamination. For a few cases, the outcome might not at all be satisfactory and sufficiently robust for a decision

Table 15.3 Scheme for supporting ERA with respect to determining the urgency of remediation at seriously contaminated sites in the Netherlands (VROM 2008, modified and currently under discussion). Depending on the land use, it is not necessary to take measures when the horizontal dimensions of the unpaved contaminated area within the contour for the Toxic Pressure (TP) is smaller than indicated. Two levels for the TP are used, i.e. $TP = 0.2$ or $TP = 0.5$

Land use	*Unpaved surface area contamination* ($TP_{MMec50}^* > 0.2$)	*Unpaved surface area contamination* ($TP_{MMec50}^* > 0.5$)
Sensitive: nature (including Ecological Main Structure and Natura 2000 areas)	$< 500\ m^2$	$< 50\ m^2$
Intermediate: agriculture, (vegetable) garden, green areas with 'nature values'**	$< 5.000\ m^2$	$< 500\ m^2$
Insensitive: other green area's, strips in build areas, infrastructure and industry	$< 50.000\ m^2$	$< 5.000\ m^2$

*TP_{MMec50} is the Toxic Pressure, which is calculated from the mixture of contaminants in soil samples (Box 15.1). The TP is calculated on the basis of total concentrations in the samples, and related to EC50 data from the literature and a mixed model for mixture effects (De Zwart and Posthuma 2005). Background concentrations of substances are subtracted from the soil sample concentrations. All concentrations are corrected for a standard soil (see Swartjes 2010). More details about the calculations can be found in Rutgers et al. (2008b).

**Outside the Ecological Main Structure and Natura 2000 areas 'nature values' are considered relevant, unless particularly stated in the petition on the land use.

about the land management regarding the contamination. In these cases, additional effort via application of the Triad approach in a subsequent tier to further improving ERA is recommended. For application of the Triad practical guidance is available (Jensen and Mesman 2006; Mesman et al. 2007; Rutgers and Den Besten 2005).

15.6.2 Examples of the Lines of Evidence in the Dutch Remediation Criterion

In the Netherlands, a practical Triad approach has been developed (Mesman et al. 2007). In Box 15.1 some examples of methods and calculation tools are presented. The three lines of evidence are described below, together with the collection of basic tools recommended for this tier of the Remediation Criterion (VROM 2008):

1. Chemistry: Model and parameters are equivalent to those at tier 2 of the Remediation Criterion (Rutgers et al. 2008b; VROM 2008). It consists of a calculation of the Toxic Pressure from the mixture of contaminants in soil samples from the contaminated site. The two-step mixed-model approach for mixture toxicity is used, combining concentration-additivity and response-additivity models (De Zwart and Posthuma 2005). In order to arrive at a number for the Toxic

Pressure which is realistic for seriously contaminated sites in the Netherlands and useful for the Triad approach, EC (Effect Concentration) data from the literature have been used (EC50 values, concentration of toxicant demonstrating 50% effect) instead of NOEC (No-Observed Effect Concentrations) data (Rutgers et al. 2008b). This is a more realistic and less conservative procedure and compatible with scaling procedures in the other lines of evidence of the Triad. A correction for bioavailability, however, is not recommended at this stage, because frameworks for implementation of bioavailability are still in development (Brand et al. 2009). Some contaminated sites, however, were assessed on the basis of testing the pore water concentrations, using basic assumptions for ERA in surface waters (Jensen and Mesman 2006; Rutgers and Den Besten 2005).

2. Bioassays: Screening level and standardized bioassays are recommended, but there is no detailed prescription. Relevant aspects for selecting a bioassay are sensitivity and validity, which should be generally accepted or carefully addressed. It is not always necessary to select bioassays with autochthonous organisms or to invest a lot of effort in this kind of tests, because differences between autochthonous and exotic organisms are usually much smaller than the differences between exposure conditions in the field and laboratory and other lab-to-field extrapolation issues (Rutgers and Den Besten 2005). Among the more popular screening tests are elutriate-based bioassays with small invertebrates, algae, plants or bacteria (e.g. *Microtox*) and whole-soil bioassay with soil invertebrates or plants. The response of the bioassays is simply expressed as a fraction of effect, ranging from 0 (no effect) to 1 (maximum theoretical effect level). More details can be found in the decision support system by Jensen and Mesman (2006).
3. Ecological observations in the field: At this stage of the Triad approach it is recommended to include plant surveys of the contaminated site and reference site(s) (with no or low contamination levels). As an alternative, simple determination of the community composition and abundance of soil invertebrates like nematodes, enchytraeids (pot-worms), earthworms and springtails may be feasible. Again, the response to the contamination should be expressed as a fraction of effect ranging from 0 (no effect) to 1 (maximum theoretical effect level). More details about the calculation of risk from multivariate observations can be found in, for example, Jensen and Mesman (2006) and Dagnino et al. (2008).

Box 15.1 Chemical characterization of effects

The Toxic Pressure (TP) from the complete mixture of contaminants in soil samples is obtained from mixture modeling using models for concentration addition (CA) and response addition (RA) (De Zwart and Posthuma 2005; Rutgers et al. 2008b). In the first steps the combination toxicity of any group with 1 or more toxicants with a comparable mode of action is calculated using the CA model:

$$\mathrm{HU_j} = \frac{[n_1]}{10^{\alpha_1}} + \frac{[n_2]}{10^{\alpha_2}} + \ldots + \frac{[n_n]}{10^{\alpha_n}} = \sum_n \frac{[n]}{10^{\alpha_n}} \text{ and: } \mathrm{TP_{CAj}} = \frac{1}{1 + e^{-\log(\mathrm{HU_j})\cdot \beta_j^{-1}}}$$

with: $\mathrm{HU_j}$ is the Hazard Units of the group of toxicants for which the CA model is valid. $[\mathrm{n_1}]$, $[\mathrm{n_2}]$, et cetera are concentrations of toxicants 1, 2,.... (e.g. in $\mathrm{mg/kg_{dw}}$; after correction for a standard soil and background concentrations; see Swartjes (2010)). α_n is a log-transformed value for toxicity (e.g. a logHC50). β_j is a slope parameter of the SSD (these α_n and β_j constants can for instance be found in Rutgers et al. 2008b).

In the next step the TP of the complete mixture of toxicants is calculated using the RA model:

$$\mathrm{TP_{MM}} = 1 - (1 - \mathrm{TP_{CA1}}) \cdot (1 - \mathrm{TP_{CA2}}) \ldots\ldots (1 - \mathrm{TP_{CAn}}) = 1 - \Pi(1 - \mathrm{TP_{CAn}})$$

The Toxic Pressure obtained from the mixed model ($\mathrm{TP_{MM}}$) is expressed as a multi substance Potentially Affected Fraction (msPAF) value and ranges from 0 (no effects) to 1 (theoretical maximum effect value).

Toxicity Characterization with Bioassays

The scaling of results from bioassays is usually straightforward, when test performance in control and reference samples is known. Sometimes, it is necessary to define a theoretical value for the full effect (1 = 100%). The final result can then be expressed as a fraction, ranging from 0 to 1. Examples of using and scaling results from bioassays can be found in e.g. Jensen and Mesman (2006) and Semenzin et al. (2008). The basic principles can be illustrated with earthworm tests. Data from the survival or reproduction of earthworms in contaminated and reference samples are straightforwardly fed into the ERA (ISO 16387:2004, ISO 11268-2:1998). The reference is set to 0; no survival is set to 1. The percentage of survival compared to the reference can be directly used as an effect value. The results from the chronic reproduction test can also follow this scheme, although arguments to use a different scale can be put forward. It becomes a bit more difficult with, for instance, the earthworm avoidance test (ISO 17512-1:2008). Typically, the distribution of worms between control and contaminated soil can be used on an effect scale (Amorim et al. 2005):

$$\text{Effect} = (R - C) \cdot (R + C)^{-1}$$

with: R is the number of worms in the reference or control soil; C is the number of worms in the contaminated soil. A negative outcome indicates attraction to the contaminated soil, which should be set to zero. Also with the avoidance test, arguments can be raised to use a different scaling, since avoidance is not straightforwardly controlling earthworm populations in the field.

Approximation of Effects from Ecological Field Monitoring

The scaling of single variables from ecological field monitoring can follow the same principles as bioassays (see above for an example). Typically, field monitoring deals with multiple variables, for which the scaling issue is less clear. The BKX-Triad algorithm to scale the results from multiple variables is very simple and robust, and can be used for virtually any dataset (Jensen and Mesman 2006), but has some drawbacks from unintended amplification in the effect calculation when many variables are present:

$$\text{BKX-Triad} = \text{Effect} = 1 - 10^{\sum_{i=1\rightarrow n} \left|\log(x_j)\cdot n^{-1}\right|}$$

With x the result of the observation i divided by the result from the reference observation and n is the number of observations at the site (or in samples).

More sophisticated scaling is possible like the use of distance values in multivariate space (e.g. Euclidean distance). This is a solution to the problem of the BKX-Triad, i.e. all deviations from the references are adding to the total calculated effect. Software tools for multi-criteria analysis are easily available, but some training is necessary to use and interpret the data:

$$\text{ED} = \sqrt{\sum\nolimits_k (y_{ki} - y_{kj})^2} \text{ and: Effect} = \text{ED}_{\text{R-C}} \cdot (\text{ED}_{\text{R-C}} + \text{ED}_{\text{C - C}theor})^{-1}$$

With ED is the Euclidian distance between site (or sample) i and j for k dimensions. Subscripts R, C and C*theor* denote the reference, contaminated and theoretically contaminated site to a 100% field effect.

After a proper scaling, the outcome of different lines of evidence should be in balance. This balance will be theoretically demonstrated with a very large number of tests addressing the set of end-points within each line of evidence. With a subset of tests, like the test proposed in the Dutch Triad approach summarized above, deviations from this balance should be expected and interpreted in terms of model uncertainties. However, if the outcome of a subset already demonstrates convergence of the results, then this is a strong basis for finalization of the ERA, providing a solid advice for the Risk Assessment or Risk Management of the site. As a practical criterion for convergence, the deviation between the outcomes of different Triad approach lines of evidence can be quantified as suggested by Jensen and Mesman (2006) and listed in Table 15.1.

15.6.3 Outline of ERA in Other Countries

The United Kingdom has developed a framework for ERA (Weeks and Comber 2005; Weeks et al. 2004). A cornerstone in this framework is the connection to the statutory regime for identification and control of sites potentially affected by contamination, also known as Part IIA of the Environmental Protection Act of 1990. This act defines a site as contaminated if:

- a contaminant source and a pathway along which the contaminant can move is present and the contaminant (potentially) can affect a specified receptor;
- there is a significant possibility of significant harm;
- contamination of controlled waters is occurring or is likely to occur.

Currently, only ecological risks to controlled waters and certain protected habitats (defined in Part IIA) are covered. The framework, however, does address how to perform ERA at sites not currently covered by Part IIA. The UK framework is based on schemes found in e.g. USA, Canada and the Netherlands. Like the procedures in these countries, it is a based on a tiered approach, where the initial tier zero aims to determine whether or not a site belongs under the Part IIA of the legislation. It involves the development of a Conceptual Site Model, which describes what is already (historically) known about the site, e.g. whether there is a likely source-pathway-receptor linkage.

In many other countries of the EU, for example Germany, Spain and Sweden, ERA can be based on additional types of testing, making a Triad approach framework feasible. A decision support system for assisting in site-specific ERA was developed based on research at the Acna di Cengio 'mega site' in the Bormida valley, Italy (Critto et al. 2007; Semenzin et al. 2007, 2008, 2009a, b).

The US and Canada were among the first in producing general frameworks for ERA (CCME 1996; US-EPA 1998). Later many amendments to the first publications were produced and these are available via the respective websites.[1] Both US and Canada frameworks for ERA address many questions related to relatively large contaminated areas, whereas some European approaches typically are designed to cope with many but smaller sites. Furthermore, wild life is a more important issue in the North America frameworks compared to the European. The reason for the somewhat reversed development of ERA in the two regions might be due to the fact that Soil Quality Standards were first developed in Europe, while general frameworks were first developed in North America. Nowadays the basic outlines of the various ERA frameworks and derivation of Soil Quality Standards world-wide seem to converge (Swartjes et al. 2008).

---

[1] http://www.epa.gov/riskassessment/ecological-risk.htm; http://www.ccme.ca/ourwork/soil.html

15.7 Outlook

In the Dutch Soil Protection Act the ecosystem is relevant for any kind of land use, although the ecological protection level varies (e.g. nature is considered more sensitive than industrial land use; Rutgers et al. (2008b); Swartjes (1999)). This has triggered attention to ERA, principally at all contaminated sites, also outside nature areas. Public support of this policy is limited, especially in cases where there is no visible damage to the terrestrial ecosystem. This has to do with lack of knowledge and, hence, appreciation for the tasks and significance of the soil ecosystem, among the general public (see Chapter 13 by Swartjes et al., this book). Results from a Triad approach-based assessment will support the acceptation of remediation measures by the general public in these cases. Acceptation is expected to improve further from increased environmental awareness due to climate change and rephrasing soil functions into the goods and services of the soil system (Rutgers et al. 2009).

A compelling recommendation to use the Triad approach in the higher tiers of ERA will result in increased attention on ecological issues and habitat protection. This was observed in the Netherlands (SKB 2009). In the lower tiers of ERA (tier one and two in the remediation criterion; Rutgers et al. (2008b)), the surface area exceeding a threshold for Toxic Pressure of the mixture of contaminants has to be remediated. In many cases in the Netherlands, the goal of a Triad approach-based ERA in tier three and subsequent tiers is then to reduce the surface area to be remediated and hence to reduce costs. Step by step, the Dutch regulators have become less hesitant with respect to interpreting Triad approach-based results. This was concluded from an inventory from 42 Triad approach-based assessments with an evaluation of the interpretation and integration of the results, and the decision-making process (SKB 2009). In 63% of the cases (total 45: the unknowns were omitted from the analysis), the management and remediation decisions were adjusted in reaction to the results from the Triad-based assessments. Since local administrations (e.g. provinces) and the national government have ratified an agreement on speeding up the soil remediation (Covenant 2009), and the procedural standard for a guidance on incorporation of a Triad in ERA will be soon available (NEN 2010), it is expected that the number of Triad-based Risk Assessments will further increase.

Although many tools for a Triad approach in ERA are available, there is still a strong demand for improved and robust methods in many cases. Also, many methods are considered not cost-effective or too laborious for smaller cases. Consequently, increasing the number of Triad-based Risk Assessments will demand for improved, new, standardized, robust and cost-effective tools.

Although ecological surveys in principle are the most site-specific part of an ERA, it is often hampered by a weak relation between contamination levels in soil and ecological observations and, hence, may lack plausibility. Ecosystems, communities and populations of organisms are shaped by a comprehensive set of environmental factors, where soil contamination is only one of those factors. Furthermore, ecological field observations occasionally need highly-trained experts and a relatively large effort.

So called ‘omic-’techniques, like genomics based on analysis of DNA-patterns, were advocated for a wider application in ERA, but progress is limited yet. The promise is that these techniques generate much and valuable information with a limited effort. However, the methods are generally immature and still quite expensive compared to traditional bioassays and community analysis. Furthermore, the issue of interpretation of community shifts for ERA is not fully resolved, i.e. scaling is still an issue. Nevertheless there is still hope for breakthroughs in this area thanks to expected technical spin-offs from medical research and agriculture business.

Further stimulating influences can be expected from international harmonization of models and frameworks for ERA by e.g. the HERACLES network (acronym: human and ecological Risk Assessment for contaminated land in European member states; Swartjes et al. (2008)). The conclusion from the network was that quite a number of Member States had readily available tools for implementing ERA. In the nearest future the Habitat and the Water Framework Directives are most likely the dominant drivers in introducing ERA in some form to a wider number of countries. In addition, also the EU soil thematic strategy and further elaborations into the future Soil Framework Directive will further stimulate attention to ERA at contaminated sites. The Triad approach will be part of these developments, triggering further improvements to previously addressed issues and new developments.

References

Amorim MJB, Römbke J, Soares AMVM (2005) Avoidance behaviour of *Enchytraeus albidus*: effects of benomyl, carbendazim, phenmedipham and different soil types. Chemosphere 59:501–510

Baird DJ, Rubach MN, Van den Brink PJ (2008) Trait-based ecological risk assessment (TERA): the new frontier? Integrated Environ Assess Manage 4:2–3

Barnthouse L (2008) The strengths of the ecological risk assessment process: linking science to decision making. Integrated Environ Assess Manage 4:299–305

Beer T (2006) Ecological risk assessment and quantitative consequence analysis. Human Ecol Risk Assess 12:51–65

Bennett JR, Kaufman CA, Koch I, Sova J, Reimer KJ (2007) Ecological risk assessment of lead contamination at rifle and pistol ranges using techniques to account for site characteristics. Sci Total Environ 374:91–101

Boivin MEY, Greve GD, Kools SAE, Van der Wurff AWG, Leeflang P, Smit E, Breure AM, Rutgers M, Van Straalen NM (2006) Discriminating between effects of metals and natural variables in terrestrial bacterial communities. Appl Soil Ecol 34:102–113

Brand E, Peijnenburg W, Groenenberg B, Vink J, Lijzen J, Ten Hulscher D, Jonker C, Römkens P, Roex E (2009) Towards implementation of bioavailability measurements in the Dutch soil regulatory framework. Report 711701084, RIVM, Bilthoven, the Netherlands

Burton GA, Batley GE, Chapman PM, Forbes VE, Smith EP, Reynoldson T, Schlekat CE, Den Besten PJ, Bailer AJ, Green AS, Dwyer RL (2002a) A weight of evidence framework for assessing sediment (or other) contamination: improving certainty in the decision-making process. Human Ecol Risk Assess 8:1675–1696

Burton GA, Chapman P, Smith EP (2002b) Weight-of-evidence approaches for assessing ecosystem impairment. Human Ecol Risk Assess 8:1657–1673

Carey JM, Burgman MA (2008) Linguistic uncertainty in qualitative risk analysis and how to minimize it. Ann NY Acad Sci 1128:13–17

CCME (1996) A framework for ecological risk assessment: general guidance. The national contaminated sites remediation program. Winnipeg, Manitoba, Canada

Chapman PM, Anderson J (2005) A decision-making framework for sediment contamination. Integrated Environ Assess Manage 1:163–173

Chapman PM, McDonald BG, Lawrence GS (2002) Weight-of-evidence issues and frameworks for sediment quality (and other) assessments. Human Ecol Risk Assess 8:1489–1515

Covenant (2009) Covenant: development of soil policy and the way to deal with remediation sites (in Dutch). Interprovinciaal Overleg, Vereniging van Nederlandse Gemeenten, Unie van Waterschappen, Ministerie van VROM en LNV. 10 July 2009

Critto A, Torresan S, Semenzin E, Giove S, Mesman M, Schouten AJ, Rutgers M, Marcomini A (2007) Development of a site-specific ecological risk assessment for contaminated sites: part I. A multi-criteria based system for the selection of ecotoxicological tests and ecological observations. Sci Total Environ 379:16–33

Dagnino A, Sforzini S, Dondero F, Fenoglio S, Bona E, Jensen J, Viarengo A (2008) A weight-of-evidence approach for the integration of environmental triad data to assess ecological risk and biological vulnerability. Integrated Environ Assess Manage 4:314–326

Danish EPA (2008) Redegørelse om jordforurening 2006. Redegørelse fra Miljøstyrelsen no 1. Danish report with English summary. Download from: http://www2.mst.dk/udgiv/publikationer/2008/978-87-7052-709-5/pdf/978-87-7052-710-1.pdf. 1 February 2008

De Zwart D, Posthuma L (2005). Complex mixture toxicity for single and multiple species: proposed methodologies. Environ Toxicol Chem 24:2665–2676

Den Besten PJ, Smidt CA, Ohm M, Ruijs MM, Van Berghem JW, Van de Guchte C (1995) Sediment quality assessment in the delta of the rivers Rhine and Meuse based on field observations, bioassays and food chain implications. J Aquatic Ecosyst Health 4:256–270

Dirven-Van Breemen EM, Lijzen JPA, Otte PF, Van Vlaardingen PLA, Spijker J, Verbruggen EMJ, Swartjes FA, Groenenberg JE, Rutgers M (2007) National land use specific reference values: a basis for maximum values in Dutch soil policy (in Dutch with English summary). Report 711701053. RIVM, Bilthoven, the Netherlands

Dirven-Van Breemen EM, Lijzen JPA, Otte PF, Van Vlaardingen PLA, Spijker J, Verbruggen EMJ, Swartjes FA, Groenenberg JE, Rutgers M (2008) National land use specific reference values: a basis for maximum values in Dutch soil policy. In: Proceedings of the 10th international UFZ-Deltares/TNO conference on soil-water systems, pp F3F6

Egler F (1977) The nature of vegetation: its management and mismanagement. Aton Forest, Norfolk, CT

Everitt BS, Dunn G (2001) Applied multivariate data analysis, 2nd edn. Arnold, London, UK

Faber JH (2006) European experiences on application of site-specific ecological risk assessment in terrestrial ecosystems. Human Ecol Risk Assess 12:39–50

Fairbrother A, Glazebrook PW, Van Straalen NM, Tarazona JV (2002) Test methods to determine hazards of sparingly soluble metal compounds in soils. Society of environmental toxicology and chemistry (SETAC), Pensacola, FL

Gaudet CL, Power EA, Milne DA, Nason TGE, Wong MP (1995) A framework for ecological risk assessment at contaminated sites in Canada. Part 1 review of existing approaches. Human Ecol Risk Assess 1:43–115

Hope BK (2006) An examination of ecological risk assessments and management practices. Environ Int 32:983–995

Hull RN, Swanson S (2006) Sequential analysis of lines of evidence – an advanced weight-of-evidence approach for ecological risk assessment. Integrated Environ Assess Manage 2: 302–311

Jensen J, Mesman M (eds) (2006) Liberation, ecological risk assessment of contaminated land, decision support for site specific investigations. ISBN 90-6960-138-9, Report 711701047, RIVM, Bilthoven, the Netherlands

Jensen J, Pedersen MB (2006) Ecological risk assessment of contaminated soil. Rev Environ Contamin Toxicol 186:73–105

Kedwards TJ, Maund SJ, Chapman PF (1999) Community level analysis of ecotoxicological field studies, 1 biological monitoring. Environ Toxicol Chem 18:149–157

Kellett BM, Beilin RI, Bristow KL, Moore G, Chiew FHS (2007) Reflecting on stakeholders' perceptions in an ecological risk assessment workshop: lessons for practitioners. Environmentalist 27:109–117

Levin R (2006) Adequacy conditions for reporting uncertainty in chemical risk assessments. Human Ecol Risk Assess 12:834–855

Linkov I, Satterstrom FK, Kiker G, Batchelor C, Bridges T, Ferguson E (2006) From comparative risk assessment to multi-criteria decision analysis and adaptive management: recent developments and applications. Environ Int 32:1072–1093

Long ER, Chapman PM (1985) A sediment quality triad: measures of sediment contamination, toxicity, and infaunal community composition in Puget Sound. Marine Pollut Bull 16:405–415

Menzie C, Hope Henning M, Cura J, Finkelstein J, Gentile J, Maughan J, Mitchell D, Petron S, Potocki B, Svirsky S, Tyler P (1996) Special report of the Massachusetts weight-of-evidence workgroup: a weight-of-evidence approach for evaluating ecological risks: report of the Massachusetts weight-of-evidence workshop. Human Ecol Risk Assess 2:277–304

Mesman M, Schouten AJ, Rutgers M, Dirven-Van Breemen EM (2007) Guideline Triad: site-specific ecological risk assessment in the remediation criterion (in Dutch with English summary). Report 711701068, RIVM, Bilthoven, the Netherlands

Natal-da-Luz T, Römbke J, Sousa JP (2008) Avoidance tests in site-specific risk assessment – influence of soil properties on the avoidance response of collembola and earthworms. Environ Toxicol Chem 27:1112–1117

Nayak TK, Kundu S (2001) Calculating and describing uncertainty in risk assessment: the Bayesian approach. Human Ecol Risk Assess 7:307–328

NEN (2010) Ecological risk assessment with regard to soil contamination (in Dutch). Norm 5737:2009 Netherlands Normalisation Institute, Delft, the Netherlands

Posthuma L, Suter GW, Traas TP (eds) (2002) Species sensitivity distributions in ecotoxicology. CRC, Boca Raton, FL

Römbke J, Eisenträger A, Hund-Rindke K, Jänsch S, Neumann-Hensel H, Schallnass (2006a) Testing and preparation of the practical use of ecotoxicological test systems (in German). Sidus, Germany

Römbke J, Sousa J-P, Schouten T, Riepert F (2006b) Monitoring of soil organisms: a set of standardized field methods proposed by ISO. Eur J Soil Biol 42: S61–S64

Rutgers M (2008) Field effects of pollutants at the community level – experimental challenges and significance of community shifts for ecosystem functioning. Sci Total Environ 405: 469–478

Rutgers M, Den Besten P (2005) Approach to legislation in a global context B. the Netherlands perspective – soils and sediments. In: Thompson KC, Wadhia K, Loibner AP (eds) Environmental toxicity testing. Blackwell, CRC, Oxford, pp 269–289

Rutgers M, Faber J, Postma J, Eijsackers H (2000) Site-specific ecological risks: a basic approach to the function-specific assessment of soil pollution. Rapporten Programma Geïntegreerd Bodemonderzoek. Volume 28, ISBN73270 448, Integrated Soil Research Programme, Wageningen, the Netherlands

Rutgers M, Mulder C, Schouten AJ, Bloem J, Bogte JJ, Breure AM, Brussaard L, De Goede RGM, Faber JH, Jagers op Akkerhuis GAJM, Keidel H, Korthals G, Smeding FW, Ter Berg C, Van Eekeren N (2008a) Soil ecosystem profiling in the Netherlands with ten references for biological soil quality. Report 607604009, RIVM, Bilthoven, the Netherlands

Rutgers M, Schouten AJ, Bloem J, Van Eekeren N, De Goede RGM, Jagers OP, Akkerhuis GAJM, Van der Wal A, Mulder C, Brussaard L, Breure AM (2009) Biological measurements in a nationwide soil monitoring network. Eur J Soil Sci 60:820–832

Rutgers M, Tuinstra J, Spijker J, Mesman M, Wintersen A, Posthuma L (2008b) Ecological risks of soil contamination in the second step of the remediation criterion (in Dutch with English summary). Report 711701072, RIVM, Bilthoven, the Netherlands

Semenzin E, Critto A, Marcomini A (2009a) Ecosystem impairment evaluation on biodiversity and functional diversity for contaminated soil. Integrated Environ Assess Manage 5(3): 399–413

Semenzin E, Critto A, Rutgers M, Marcomini A (2007) Development of a site-specific ecological risk assessment for contaminated sites: Part II. A multi-criteria based system for the selection of bioavailability assessment tools. Sci Total Environ 379:34–45

Semenzin E, Critto A, Rutgers M, Marcomini A (2008) Integration of bioavailability, ecology, and ecotoxicology by three lines of evidence into ecological risk indexes for contaminated soil assessment. Sci Total Environ 389:71–86

Semenzin E, Critto A, Rutgers M, Marcomini A (2009b) DSS-ERAMANIA: Decision support systems for site-specific ecological risk assessment of contaminated sites. In: Marcomini A, Suter GW II, Critto A (eds) Decision support systems for risk based management of contaminated sites. Springer, Germany

Sijm DTHM, Van Wezel AP, Crommentuijn T (2002) Environmental risk limits in the Netherlands. In: Posthuma L, Suter GW, Traas TP (eds) Species sensitivity distributions in ecotoxicology. CRC, Boca Raton, FL, pp 221–254

SKB (2009) Evaluation of the application of the Triad approach for ecological risk assessment (in Dutch). Stichting Kennisontwikkeling en Kennisoverdracht Bodem (SKB), project report PTS 808, Gouda, the Netherlands

Suter GW, Efroymson RA, Sample BE, Jones DS (2000) Ecological risk assessment for contaminated sites. CRC, Lewis, Boca Raton

Swartjes FA (1999) Risk-based assessment of soil and groundwater quality in the Netherlands: standards and remediation urgency. Risk Anal 19:1235–1249

Swartjes FA, Carlon C, De Wit NHSM (2008) The possibilities for the EU-wide use of similar ecological risk-based soil contamination assessment tools. Sci Total Environ 406:523–529

US-EPA (1998) Guidelines for ecological risk assessment. EPA/630/R-95/002F. Risk Assessment Forum, Washington, DC

Versluijs CW, Mulder HFMM, Van Wijnen HJ, Van den Broek HH (2007) Soil remediation in areas with political priorities (in Dutch with English summary). Report 607700001. RIVM, Bilthoven, the Netherlands

VROM (2008) Circulaire Bodemsanering 2006. Staatscourant 10 Juli 2008, 131/p 23; 1–28. SDU, The Hague, the Netherlands

Walker WE, Harremoes P, Rotmans J, Van der Sluijs JP, Van Asselt MBA, Janssen P, Krayer von Krauss MP (2003) Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. Integrated Assess 4:5–17

Weed DL (2005) Weight of evidence: a review of concept and methods. Risk Anal 25:1545–1557

Weeks JM, Comber SDW (2005) Ecological risk assessment of contaminated soil. Mineral Mag 69:601–613

Weeks J, Sorokin N, Johnson IJ, Whitehouse P, Ashton D, Spurgeon D, Hankard P, Svendsen C (2004) Biological test methods for assessing contaminated land. Stage 2 – a demonstration of the use of a framework for the ecological risk assessment of land contamination. Science group report P5-069/TR1. ISBN: 1844322963. Environment Agency, UK

Wright DA, Welbourn P (2002) Environmental toxicology. University, Cambridge, UK

Chapter 16
Bioavalibility in Soils

Mark E. Hodson, Martina G. Vijver, and Willie J.G.M. Peijnenburg

Abstract In this chapter we review and discuss the commonly used phrase or concept "bioavailability". This concept is key to Risk Assessment as it assesses what proportion of a contaminant present at a contaminated site is available for uptake by organisms and is thus potentially able to cause harm. Whilst this is a relatively straightforward concept the reader will discover that in reality life is not that simple. We start by reviewing the different definitions of bioavailability currently in use. We go on to discuss how soil properties impact on the bioavailability of both metal, metalloid and organic contaminants. Next we review the different methods people currently use to determine bioavailability, concentrating on chemical extractions, but also covering modelling approaches. We conclude that a precise definition of bioavailability equally applicable to all different contaminated sites, contaminants and organisms is unlikely to be achieved. Similarly, a single chemical extraction is unlikely to give a universal measure of bioavailability. However, the message is not all doom and gloom. On a contaminant by contaminant or species by species level chemical extractions and other measurement techniques can accurately predict bioavailability. Modelling techniques are constantly improving and offer hope for the future in terms of predicting bioavailability. At present however, the best method of determining the amount of contaminant available for uptake by an organism is to measure the concentration of the contaminant in the organism. Even this method, however, is open to question as organisms can and have evolved methods of regulating metal uptake.

M.E. Hodson (✉)
University of Reading, Reading, UK
e-mail: m.e.hodson@reading.ac.uk

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_16, © Springer Science+Business Media B.V. 2011

Contents

16.1 Introduction

The initial measurements of contaminants during site investigation are usually total concentrations of contaminants. However, it is usually the "bioavailable" fraction that is relevant to whether the contaminants present in a soil pose a hazard or not. A search on the Internet based search engine Web of Knowledge for peer-reviewed journal articles concerned with "Bioavailability" and "Soil" returns 7703 entries with a near perfect linear increase in journal articles for over a decade ($y = 55x - 109969$, $p \leq 0.001$ as of August 2008!). Add to this the vast number of government and agency reports that exist on the subject and the reader will appreciate that this chapter can not be fully exhaustive on its subject! Given the large amount of research effort and funding that has been expended on the subject it is perhaps surprising that there is neither a universally accepted definition of bioavailability nor a standardized means of measuring it. This deficiency reflects in part the complexity and diversity of the soil system and in part the legislative requirements of different jurisdictions. This chapter aims to summarise for organic and inorganic substances: (1) current thoughts on what bioavailability is, (2) how soil properties impact bioavailability and, (3) methods for measuring and predicting bioavailability. Covered elsewhere in this book, but of relevance to bioavailability, are chapters on oral bioavailability to humans (see Chapter 7 by Wragg et al., this book), uptake of metals from soil into vegetables (see Chapter 8 by McLaughlin et al., this book) and uptake of organic contaminants from soil into vegetables (see Chapter 7 by Trapp and Legind, this book). Besides, bioavailability is relevant for site specific ecological Risk Assessment (see Chapter 15 by Rutgers and Jensen, this book).

From the perspective of contaminated sites and their remediation the important question is not "How bioavailable is a particular element?" but, "Will this contaminant have a toxic effect at this soil concentration on this organism or other receptor?" Clearly bioavailability is an important part of this question but for the assessment of contaminated sites bioavailability needs to be tied to a consideration of toxic effects or legislatively permissible concentrations of contaminants in selected target organisms.

16.2 What is Bioavailability?

In the last five years a variety of international working parties and governmental organisations have produced reports on bioavailability. There has also been a healthy discussion in the scientific literature regarding the term. The emerging consensus appears to be that, given the variability present in the environment, in terms of soil properties and soil organisms a definition of bioavailability that is precise, universal and useful is unlikely to be attained. However, there is an emerging consensus on what the term bioavailability encompasses, how it should be used and what information should be provided on a case by case basis to make the term meaningful and useful.

In 2002 the United States of America National Research Council published a report on bioavailability in soils (National Research Council 2002) which was usefully summarised by Ehlers and Luthy (2003). The report notes that the term bioavailability has been defined by various disciplines and goes on to state that this has led to some confusion over the term and that therefore further definitions will be avoided. Instead the report focuses on *bioavailability processes*, which are the biological, chemical and physical processes that result in an organism being exposed to a contaminant present in the soil. These processes are: release of the contaminant from the solid phase, transport of the contaminant to and across a biological membrane and, incorporation into a living organism. The National Research Council report goes on to document methods for measuring these processes. Partially in response to the National Research Council report, but also noting the legal and regulatory implications of the concept of bioavailability within the Risk Assessment process, Semple et al. (2004) proposed the introduction of the term “bioaccessibility” to complement bioavailability. According to Semple et al. bioavailable material is “freely available to cross an organism’s cellular membrane from the medium the organism inhabits at a given time”, whilst bioaccessible material is “that which is available to cross an organism’s cellular membrane from the environment, if the organism has access to the contaminant”. The key advance offered by Semple et al’s definitions is the inclusion of a time element. Within their paper Semple et al. note that many classical chemical extraction methods for measuring “bioavailability” (see Section 16.4.1) actually measure or predict bioaccessibility and that, for Risk Assessment and remediation, it is often bioaccessibility rather than bioavailability that is of concern. Semple et al. also make the points that bioavailability and bioaccessibility will vary between organisms and that the “membrane” in question can vary with the organism, for example being just the cellular membrane for a bacterium, but including both the epidermis and gastrointestinal tract of an earthworm. Following Semple et al. (2004) the term “bioaccessibility” is gaining usage in the literature. Reichenberg and Mayer (2006) made a further contribution to the debate over what bioavailability is when they proposed that bioavailability comprised both “accessibility” and “chemical activity”. Their “accessibility” term is essentially the same as the Semple et al. (2004) bioaccessibility term. Chemical activity is a classically defined term related to fugacity and freely dissolved concentrations, which quantifies the partial molar free energy of a substance (Faure 1992). A variety of *bioavailability processes* are driven by chemical activity, for example, sorption, diffusion and partitioning occur from high to low activities. The use of chemical activity as a component of bioavailability had already been incorporated into such models as the *biotic ligand model* (e.g. Di Toro et al. 2001).

The International Standards Organisation (ISO) has established a working group on bioavailability (ISO/TC190 – Soil Quality). Harmsen (2007) summarises their activities. The definition of bioavailability used by this organisation is "Bioavailability is the degree to which contaminants present in the soil may be absorbed or metabolised by human or ecological receptors or are available for interaction with biological systems" (ISO 2005) which is similar to the definition of bioaccessibility given by Semple et al. However, Harmsen (2007) argues that bioavailability is best presented as a concept that is related to specific situations or measurements on a case by case basis.

To conclude, a variety of definitions for bioavailability have been proposed in the literature. There is a growing consensus that in broad terms it is a measure of uptake of and consequent cellular interaction with a contaminant by an organism. There is also a growing acknowledgement that bioavailability will: (1) vary between organisms, (2) be a function of time of contact between the organism and contaminated media and, (3) that further, more precise definitions will only be applicable on a case by case basis. The rest of this chapter will review the impacts that soil properties have on bioavailability and the various methods (bioassays, chemical extractions and models) that exist to measure or predict bioavailability.

16.3 Impact of Soil Properties on Bioavailability

To be bioavailable molecules must cross a biological membrane. In effect this means that the molecules have to interact with the aqueous phase, either the soil's pore water or in soil passing through the gastrointestinal tract of an organism. Therefore, soil properties which control partitioning between the solid phase in soil and the pore water, such as pH, organic matter content, Eh, cation exchange capacity (CEC), and the concentration of clay minerals and oxyhydroxides, have a significant impact on bioavailability. Generally, molecules held on exchange sites can readily become bioavailable by desorption from these sites. Indeed many studies treat molecules held on exchange sites as part of the bioavailable fraction despite the contradiction that this generates with the above definitions of the term. As a consequence, an increase in the number of exchange sites aids the retention of molecules in a bioavailable form. Molecules sorbed more strongly to surfaces or in solid form are generally not available, whilst molecules in the pore water are available, but are also prone to being removed from the soil system as solution leaches through the soil.

From here on, inorganic and organic contaminants will be discussed separately, because their chemical and ecotoxicological properties are rather different. This results in different fate and consequently different interactions with living cells. The main characteristics of each group are summarized in Table 16.1 (taken from Campbell et al. 2006).

Table 16.1 Summary of the general characteristics that determine the kinetics of uptake and depuration of organic contaminants, metals, and inorganic metal contaminants in biota

Organics	Metals and metalloids
Tissue uptake is most commonly a blood flow limited process, with linear partitioning into tissues.	Metals and their complexes are often ionized, with *tissue uptake* (membrane transport), having greater potential to be diffusion-limited or to use specialized transport processes.
Metabolism is generally extensive and often species specific.	*Metabolism* is usually limited to oxidation state transitions and alkylation/de-alkylation reactions.
Persistence in body fat is common, because of lipid solubility (not capacity-limited).	Often *sequestered*, bound to specific plasma or tissue proteins (intrinsically capacity-limited), or deposited in inert forms like waste nodules, mineral concretions and granules.
May be *eliminated* by excretion in urine after biotransformation from lipophilic forms to hydrophilic forms, due to complex metabolism.	Predominantly *eliminated* by excretion because metal contaminants are generally small molecules and are hydrophilic. Thereupon excretion of inert species.
Generally substance-specific *homeostatic mechanisms*[a] are not available.	Essential metals have *homeostatic mechanisms* that maintain optimum tissue concentrations over a range of exposure rates.
Interactions with other structurally similar contaminants may occur, especially during metabolism.	*Interactions* among metals and between metals and organics are numerous and occur commonly during the processes of absorption, excretion, and sequestration.

Taken from Campbell et al. (2006)

[a] Homeostatic mechanisms are physiological mechanisms by which organisms maintain concentrations of essential elements inside their body within specific fixed levels. These fixes levels are often termed the “window of essentiality”

16.3.1 Metals and Metalloids

For inorganic contaminants present as cations the bioavailability tends to decrease from acid to neutral through alkaline conditions. Typically, metals are more soluble under acidic conditions. However, over time, a highly acid soil is likely to loose solution cations through leaching. As a consequence, in the longer term the concentration of bioavailable contaminants can be reduced by acidic conditions, unless there is a significant source of contaminants to replenish those lost, for example the gradual dissolution of metal-bearing materials. Loss of cations through leaching can be offset by exchange processes. Cations held on exchange sites are normally considered to readily become bioavailable. As pH increases from acid to alkali, metal ions are more likely to displace protons from exchange sites and be held in the soil in a form that allows rapid replenishment of the reservoir of bioavailable contaminants as they are taken up by organisms. Lindsay (1979) presents a series of empirical equilibrium constants and phase diagrams for the reaction:

$$\text{Soil-metal} + \text{xH}^{+} \Leftrightarrow \text{soil-H} + \text{metal}^{x+} \tag{16.1}$$

Soil-metal and soil-H indicate metal ions and protons sorbed to exchange sites H^+ and metal^{x+} represent protons and metal ions in solution.

Besides, Lindsay (1979) gives partition co-efficients and solubility data for a range of metals, that allow predictions of conditions under which specific metal ions are likely to reside in the pore water or on exchange sites in soils.

In contrast to cationic contaminants, contaminants present as anions such as arsenate and chromate tend to be more significantly sorbed at lower pH, thus becoming more bioavailable (but also more mobile in the soil) towards more neutral pH conditions.

pH influences bioavailability via the generation of exchange sites as well as exerting controls via solubility and competition for exchange sites between protons and other cations. The charge of mineral surfaces may be permanent or pH-dependent. The former, i.e. the permanent charge of minerals, results from structural imperfections in mineral structures. The latter, i.e. the pH-dependent charge of minerals, results from dissociation (leading to negative surface charge) or protonation (leading to positive surface charge) of hydroxyls associated with the surface of clay minerals, oxyhydroxides or organic matter. The relative importance of permanent and pH-dependent charge of minerals varies between metals. Generally, permanent charge is most important for so-called *2:1 clays* such as vermiculite, which comprise a basic repeating structure of Al-octahedral layers with a Si-tetrahedral layer sandwiched in between. pH dependent charge is more important for *1:1 clays* such as kaolinite (i.e. a clay with alternating layers of Al-octahedral and Si-tetrahedral sheets) and oxyhydroxides. The pH at which the surface charge switches from positive to negative (the point of zero charge, PZC) varies with sorbent, being typically 7–9 for Fe-oxyhydroxides, 8–9.2 for Al-oxyhydroxides, 1.5–4.6 for Mn-oxyhydroxides and 5–6 for clay minerals. These PZC values mean that in “typical” soils with a slightly acidic pH, Fe- and Al-oxyhydroxides are important for retaining bioavailable anions in the soil and Mn oxyhydroxides for retaining cations.

Surfaces of organic matter can have a negative charge through the dissociation of carboxylic acid and phenolic acid groups and thus can provide exchange sites for cationic metals. However, metals and organic matter also interact to form chelate complexes in which the metal is sorbed to the organic matter through more than one bond so that a ring structure is formed. As an illustration of the formation of chelate complexes, cadmium is complexed with one deprotonated carboxylate group (Fig. 16.1a), and copper is complexed with one carboxylate and one neighbouring phenolate group (Fig. 16.1b):

The resulting ring structures are very stable. There is an element of covalency in the bonds and thus metals held in this way are not readily bioavailable. Thus, increasing organic matter often leads to a decrease in the bioavailability of metals. Different metal cations show different tendencies to form complexes with organic

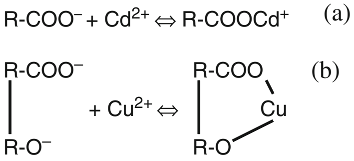


Fig. 16.1 Cadmium complexed to one deprotonated carboxylate group (**a**), and copper complexed to one carboxylate and one neighbouring phenolate group, as an illustration of the formation of chelate complexes

matter (McBride 1994), with copper in particular able to form a strong bound. Overall, total metal-binding concentrations to organic matter decrease in the order copper > nickel > lead > cobalt > cadmium > calcium > zinc > manganese > magnesium (McBride 1994). Unfortunately, there is currently no systematic explanation for this. The difficulty in determining the reason for this variation is probably due to the wide variety of functional groups present on organic matter. In the "real" world, the strength of metal binding is to a considerable extent dependent on competition for binding sites with other commonly occurring cations, such as H^+ and macro-ions like Ca^{2+}, Mg^{2+}, Na^+, et cetera.

The redox conditions in soil will also impact bioavailability of metals. In general terms, for elements that can exist in more than one oxidation state the lower oxidation state ions are more soluble. So under more reducing conditions the concentration in the pore water often increases. If soils are water-logged and become anaerobic, oxyhydroxides of Fe and Mn become unstable and dissolve. Any sorbed ions are released and there is an initial increase in bioavailability. Over time, bioavailability can decrease as pore water is leached from the upper soils. In waterlogged soils, however, leaching rates are relatively slow. Release of ions into the pore water initiated by waterlogging is partially offset by the precipitation of sulphides, which can reduce metal availability. Indeed, some in situ remediation techniques rely on generating reducing conditions to render metals unavailable.

Finally, there is increasing evidence that speciation of metals governs their bioavailability. Indeed, models such as the *Biotic ligand model* (Section 16.4.2.1) rely on this phenomenon (Allen et al. 2008; Arnold et al. 2007; Di Toro et al. 2001; Lock et al. 2007; Steenbergen et al. 2005; Thakali et al. 2006a, b; Van Gestel and Koolhaas 2004). Generally, metals present as free ions or as simple, relatively small inorganic complexes are viewed as being available, whilst larger, often organic, complexes are not. In general, the higher the ionic strength of the pore water, the less free ions are present, and therefore the lower the bioavailability of that metal.

16.3.2 Organic Contaminants

The factors affecting the bioavailability of organic contaminants are fundamentally the same as those for metals, but the magnitude of effects varies. To assess the bioavailability of organic contaminants it is also necessary to determine the physical properties of the contaminants and to understand their biogeochemical fate, requiring the need to investigate transport properties, sorption, volatilization, and biodegradation of organic contaminants. The more of the contaminant present in the pore water the more bioavailable it is. A popular aphorism used for predicting solubility is "*like dissolves like*" (Williamson et al. 2007). This indicates that a contaminant will dissolve best in a solvent of similar polarity. This is a rather simplistic view, since it ignores many solvent-solute interactions, but it is a useful rule-of-thumb. The behaviour of organic contaminants is often characterized by octanol-water partitioning, expressed as the octanol-water partitioning coefficient ($\mathrm{K_{ow}}$). The larger the $\mathrm{K_{ow}}$, the stronger the contaminants are bound to the soil solid phase and the less bioavailable they are. In general, the hydrophobicity and hence the $\mathrm{K_{ow}}$ increase with increasing number of C-atoms.

The solubility of organic contaminants nearly always increases with temperature. The solubility equilibrium is relatively straightforward for non-ionic and non-polar substances such as benzene. When dissolved in water, the benzene molecules remain intact and are generally surrounded by water molecules. This behavior differs from that of ionic contaminants or metal salts. When an ionic contaminant such as sodium chloride (NaCl) dissolves in water, the sodium chloride lattice dissociates into individual ions that are solvated or surrounded by water molecules, thus increasing the ionic strength of the solution. In turn, increases of ionic strength affect the surrounding by water of neutral organic contaminants. Effectively, at high ionic strength conditions, the solubility of organic contaminants and hence their bioavailability is reduced.

Apart from temperature and ionic strength, sorption to (dissolved) organic matter is the main factor affecting bioavailable concentrations of organic contaminants. Sorption increases with increasing organic matter content of the soil, whereas concentrations of organics in the pore water increase with increasing dissolved organic carbon (DOC) concentrations. In contrast to a large number of aquatic organisms, for which only the truly dissolved contaminant has been shown to be bioavailable, it can, however, not be ruled out that contaminants sorbed onto dissolved organic carbon can interact with biota. Amongst others, this is due to differences in the rates of release of organics from dissolved organic carbon versus release kinetics from soil solid organic material. Karickhoff et al. (1979) were one of the first authors to show the equilibrium concept of partitioning of organic contaminants by reporting that the contaminant-specific organic-carbon normalized partition coefficient (K_{oc}) is proportional to K_{ow}:

$$(\mathrm{Log})K_{oc} \approx \mathrm{a} * (\mathrm{Log})K_{ow} + \mathrm{b} \tag{16.2}$$

where a and b are contaminant specific constants.

Subsequently, the K_{oc} may be used to predict the degree of contaminant partitioning of hydrophobic organics between soil organic carbon and pore water:

$$K_{oc} = K_d / f_{oc}, \text{with}\, K_d = C_w / C_{\text{solid phase}} \tag{16.3}$$

where f_{oc} is the fractional organic matter content of the soil and C_w and $C_{\text{solid phase}}$ are the concentrations of the contaminant in the pore water and soil solid phase, respectively.

Based on additional work of Sabljić et al. (1995), 19 quantitative structure-activity relationships (QSARs) for a variety of contaminant classes are listed in the EU Technical Guidance Document (TGD – European Commission 2003) for Risk Assessment of chemicals in the form of Eq. (16.2) (Table 16.2). Low polarity organic contaminants, however, may, on top of equilibrium partitioning, also bind to solid phases through adsorption mechanisms which result in greater (non-linear) binding coefficients. Thermally or diagenetically altered forms of carbonaceous materials such as coal, kerogen from shales, soot, and charcoal, have a particularly high binding affinity and nonlinear adsorption behaviour, with carbon-normalized Freundlich sorption coefficients that are as much as 50–250 times higher than typically reported K_{oc} values (Binger et al. 1999; Bucheli and Gustaffson 2000; Grathwohl 1990).

Hydrophobic partitioning is less important for polar and ionisable contaminants. These contaminants are involved in more diverse binding mechanisms that contribute to contaminant retention, including ion bonding or ligand exchange, covalent binding to the soil molecular structure, ion-dipole and dipole-dipole interactions, charge transfer, hydrogen bonding and hydrophobic bonding (Van der Waals forces; Von Oepen et al. 1991). The chemically most active component of the soil is the colloidal fraction which consists of organic matter and inorganic clay minerals (Stevenson 1994). Both components display a negative electrical charge at the surface, resulting in weak binding forces and typically reversible sorption, as sorption often is restricted to a limited number of binding sites at the surface layer. The effect of the negative electrical charge at the surface can be measured by the cationic exchange capacity, which on average is 50 meq/100 g for clays and 290 meq/100 g for humic acids (Krogh 2000). Electrical forces involving charge transfer (~40 kJ/mol) are stronger than hydrophobic bonding (~4 kJ/mol) (Von Oepen et al. 1991) so that they dominate when present. Thus, a different degree of sorption of anions, cations and neutral molecules can be expected, with cations showing the highest potential for sorption, due to electrical attraction. Although lipophilic interactions are weaker than the other interactions mentioned, they are the most important for the majority of organic contaminants.

A typical example of this general sorption behaviour is provided by Vasudevan et al. (2002). These authors studied the sorption of two ionic contaminants (2,4-D and quinmerac) and one neutral chemical (norflurazon) onto iron oxide-rich, variable charged soils. It was found that sorption of 2,4-D and quinmerac was strongly influenced by mineralogy, particularly soil iron and aluminium oxides, whereas sorption of the neutral norflurazon was only related to total soil carbon. An appreciable fraction of the mass sorbed in stirred-flow studies was easily desorbed by deionised water, and desorption of ionic contaminants was initially more rapid than sorption. This sorption-desorption behaviour, although contrary to desorption hysteresis commonly observed in batch studies, suggests that the reversibly sorbed fraction is weakly bound to the soil surface. 2,4-D sorption to iron oxide-rich

Table 16.2 Overview of quantitative structure-activity relationships (QSARs) (K_{oc}/K_{ow} relationships) for a variety of contaminant classes, included in the TGD

Contaminant class	Equation
Predominantly hydrophobics	$K_{oc} = \frac{1.26 \cdot K_{ow}^{0.81}}{1000}$
Nonhydrophobics	$K_{oc} = \frac{10.47 \cdot K_{ow}^{0.52}}{1000}$
Phenols, anilines, benzonitriles, nitrobenzenes	$K_{oc} = \frac{7.94 \cdot K_{ow}^{0.63}}{1000}$
Acetanilides, carbamates, esters, phenylureas, phosphates, triazines, triazoles, uracils	$K_{oc} = \frac{12.30 \cdot K_{ow}^{0.47}}{1000}$
Alcohols, organic acids	$K_{oc} = \frac{3.16 \cdot K_{ow}^{0.47}}{1000}$
Acetanilides	$K_{oc} = \frac{13.18 \cdot K_{ow}^{0.40}}{1000}$
Alcohols	$K_{oc} = \frac{3.16 \cdot K_{ow}^{0.39}}{1000}$
Amides	$K_{oc} = \frac{17.78 \cdot K_{ow}^{0.33}}{1000}$
Anilines	$K_{oc} = \frac{7.08 \cdot K_{ow}^{0.62}}{1000}$
Carbamates	$K_{oc} = \frac{13.80 \cdot K_{ow}^{0.37}}{1000}$
Dinitroanilines	$K_{oc} = \frac{83.18 \cdot K_{ow}^{0.38}}{1000}$
Esters	$K_{oc} = \frac{11.22 \cdot K_{ow}^{0.49}}{1000}$
Nitrobenzenes	$K_{oc} = \frac{3.55 \cdot K_{ow}^{0.77}}{1000}$
Organic acids	$K_{oc} = \frac{2.09 \cdot K_{ow}^{0.60}}{1000}$
Phenols, benzonitriles	$K_{oc} = \frac{12.09 \cdot K_{ow}^{0.57}}{1000}$
Phenylureas	$K_{oc} = \frac{11.22 \cdot K_{ow}^{0.49}}{1000}$
Phosphates	$K_{oc} = \frac{14.79 \cdot K_{ow}^{0.49}}{1000}$
Triazines	$K_{oc} = \frac{31.62 \cdot K_{ow}^{0.30}}{1000}$
Triazoles	$K_{oc} = \frac{25.70 \cdot K_{ow}^{0.47}}{1000}$

European Commission (2003)

soils and pure-phase metal oxides appears to be driven by non-specific electrostatic attraction, with specific electrostatic attraction and van der Waals interactions being secondary. Both the carboxylate and the heterocyclic nitrogen groups may participate in sorption of quinmerac, facilitated by specific and non-specific electrostatic attraction with charged soil mineral surfaces, and surface complexation. The heterocyclic nitrogen, amine, and carbonyl groups of norflurazon do not appear to interact with soil minerals.

A phenomenon that is observed when studying sorption of lipophilic, polar and ionic contaminants is related to the displacement of contaminants initially sorbed to the solid phase. This is for instance reported for sorption of pharmaceuticals to sewage sludge by Jones et al. (2006). This phenomenon may be ascribed to initially fast binding of the contaminant to a large number of low affinity binding sites (like weak binding to the surface layer), followed by displacement of the contaminants by for instance more hydrophobic constituents of the DOC or by (newly formed) DOC constituents capable of interacting more strongly with the sorption sites. Thereupon, these constituents may out-compete micro contaminants as they may be formed in relative high concentrations due to (microbial) degradation of the organic matter present in either the solid phase or in the pore water.

Bintein and Devillers (1994) suggested an empirical regression using K_{ow} and pK_a as molecule descriptors and pH and f_{OC} to describe the sorbent properties for the estimation of the K_d of both ionised and non-ionised contaminants:

$$\log K_d = 0.93 \cdot \log K_{OW} + 1.09 \log f_{OC} + 0.32 \cdot CF_a - 0.55 \cdot CF_b + 0.25 \quad (16.4)$$

where the correction factors CF_a and CF_b quantify the fraction of dissociated acids and bases in the system:

$$\begin{aligned} CF_a &= \log \frac{1}{1 + 10^{pH - pKa}} \\ CF_b &= \log \frac{1}{1 + 10^{pKa - (pH-2)}} \end{aligned} \quad (16.5)$$

According to these authors, the soil pH should be entered for acids into the equation for the correction factor CF_a. The pH required for the correction factor CF_b is the pH at the surface of the soil colloids and is lowered 2 pH units below the pH of the bulk soil solution. Even though the coefficient of determination of this equation was very high ($R^2 = 0.93$), a short-coming of the regression, regarding charged organic contaminants, is their insufficient representation in the test set (nine acids and three bases out of 87 contaminants). Besides, it may also surprise that the fraction of bases, CF_b, was negatively correlated to the sorption coefficient K_d, even though it might be expected that electrical attraction of the positively charged cations by the negative potential of soil colloids would increase sorption of bases.

In summary, although the predictability of newly developed empirical formulae is large, there is no rigorously tested method available to predict the K_d or the K_{oc} of the majority of contaminants. Separate expressions are needed for the neutral and the ionic molecule fraction in order to account for dissociation of organic acids, bases and amphoters.

16.4 Measurement of Bioavailability

The simplest and most straightforward method of determining the bioavailable fraction of a contaminant in a soil is to expose the organism of interest, be it an earthworm or a geranium, and measure the uptake of the contaminant into that organism. In the case of organic contaminants the bioavailable fraction could be assessed by measuring the rate of biodegradation, i.e. perform a bioassay. This gives a direct measure of what is bioavailable to a specific organism over a specific period of time. Clearly this is not practical as there are too many contaminated sites, too many potential receptors and too many contaminants for direct bioassays to be performed as a matter of standard practise. For this reason, a variety of methods have been proposed to determine bioavailability. These can be divided into models that quantify scenarios on bioavailability and chemical extractions that mimic bioavailability and subsequent uptake. Because of the differences in their behaviour (Table 16.1), measurement procedures vary for inorganic and organic contaminants. Measurement procedures which are the same for the two groups of contaminants are the physical sampling of pore water (see MacDonald et al. (2008) for a review of methods) allied to the assumption that contaminants in the pore water are bioavailable. Methods for sampling pore water include simple centrifugation of the soil (e.g., Nahmani et al. 2007a), vacuum sampling (e.g., Tipping et al. 2003) and passive sampling systems based on a zero-tension principle (e.g. Haines et al. 1982). The most common vacuum sampler is the suction cell, where an imposed vacuum sucks pore water into an in-ground porous cup. In simple systems, the water is stored in the suction cell and is subsequently sucked or blown into a sample flask placed on the soil surface. When employing these methods in the laboratory it is common to standardize the soil in advance. The most common standardization treatment is fixing the water content in the soil and allowing the system to equilibrate for a specific period of time (e.g. Tipping et al. 2003). Water concentrations are commonly expressed in terms of percent of the maximum water holding capacity.

16.4.1 Extractions for Determining Bioavailability

Chemical extractions represent perhaps the most straightforward method for determining bioavailability of a contaminant. These methods typically involve shaking contaminated soil material with a solution for a period of time and then analysing the concentration of contaminants in the solution. Contaminants that are more readily extracted are more bioavailable. For standard extractions the ratio of solid to solution, period of shaking, nature of shaking, temperature, et cetera, are stipulated in guidelines. Any chemical extraction of this nature should have been validated against field based measurements, showing good correlations between extractable chemicals and tissue concentrations measured in an organism of interest. A problem is that the extractions are often applied outside the limits of their validation, e.g., different soil texture, pH, contaminant concentration, organic matter concentration.

The majority of chemical extractions have been developed for plants. There are some attempts in the literature to apply these to soil biota, e.g. earthworms (Conder et al. 2001; Langdon et al. 2001), but these are rather limited and are currently used far less frequently than plant correlated extractions.

Chemical extractions as a proxy for bioavailability are popular due to their rapidity and cost effectiveness. Provided that the limitations of such methods are borne in mind, and that results are presented as extractable total concentrations rather than bioavailable concentrations, such extractions are useful tools in Risk Assessment. However, there will always be a problem with extractions. Due to the infinite variability between organisms, soils and different elements, it will always be hard to have a high degree of confidence that a particular extractable concentration is a direct measure of bioavailability for a specific organism, especially under changing conditions. For this reason, calculating procedures perhaps offers a better solution to measure the bioavailability in the long term.

16.4.1.1 Metals and Metalloids

Chemical extractants used to assess the bioavailability of inorganic contaminants involve solutions of either simple ionic solids, e.g., $CaCl_2$, or more complex organic reagents such as ethylenediaminetetraacetic (EDTA). The simple ionic solutions are used as a proxy for pore water. They are used in order to (1) dissolve readily soluble contaminants and (2) displace contaminants from exchange sites into solution. The more complex organic reagents are supposed to mimic the organic exudates produced by plants, which are able to more aggressively remove contaminants from exchange sites into solution. There are too many different extractions to review them all in this chapter. Different extractions are well documented in, amongst others, Dean (2007), McLaughlin et al. (2000), Houba et al. (1996) and Ross (1994). Perhaps the most commonly used chemical extractants to assess the bioavailability of metals are 0.01 M $CaCl_2$, diethylenetriamenepentaacetate (DTPA) and EDTA. They are described briefly below.

The first suggestion that 0.01 M $CaCl_2$ extractions could be used as a general assay for bioavailability that the authors are aware of was made by Houba et al. (1996). They suggest that < 2 mm air-dry soil is extracted at a ratio of 10 mL of 0.01 M $CaCl_2$ to 1 g of soil by shaking at 20°C for 2 h, after which the pH is measured, the mixture centrifuged and the centrifugate analysed for the elements of interest. Advantages of the 0.01 M $CaCl_2$ extraction procedure are that: the concentration of the solution is in the range of that of "average" soils, the Ca^{2+} causes clays to coagulate, as Ca is the dominant adsorbed cation on soil exchange sites $CaCl_2$ solution is better able to displace metals from exchange sites than other extractants, minimal chemicals are used reducing extractant disposal issues, and a single extractant can be used for a multitude of metals. Houba et al. (1996) are careful to avoid attributing all $CaCl_2$ extractable metals to an adsorbed fraction, recommending instead, expressing extractable metal concentrations in terms of solute concentrations in the extractant pore water concentrations. A variety of studies were cited by Houba et al. (1996) in which 0.01 M $CaCl_2$ extractable metals have been correlated with metal concentrations in plants and many studies have been carried out since (e.g., Brun et al. 2001; Fang et al. 2007; Mench et al. 1997; Perez-de-Mora et al. 2006). A slightly different approach is to use $CaCl_2$ as the extractant but, in an effort to mimic pore water conditions, to tailor the concentration of solution used on a soil by soil basis. Thus, to simulate the metal concentrations in the pore water of a soil, the first step would be to determine the ionic strength of the pore water. For example, for clayish soils ionic strengths are typically around 2.5 mM (Schröder et al. 2005), implying that 0.01 M $CaCl_2$ will overpredict the pore water concentrations. As indicated by Schröder et al., 2.5 mM $CaCl_2$ as extractant would be a better choice in this example to mimic pore water.

DTPA extractions are perhaps the most used and abused procedures to measure bioavailability. Lindsay and Norvell (1978) proposed a method (shake 10 g $<$ 2 mm air dry soil at a rate of 120 cycles min^{-1} on a horizontal shaker with a stroke of 8 cm in a solution comprising 0.005 M DTPA, 0.01 M $CaCl_2$ and 0.1 M TEA buffered to pH 7.30 using HCl, then filter through Whatman No. 42 filter paper and analyse) for determining deficiencies of iron, manganese, zinc and copper in soils. The method was adopted as a mean to assess plant bioavailable metals at contaminated sites. O'Connor (1988) documents the main problems with using this method for soils containing high concentrations of metals, the most significant being that the chelating capacity of the solution (10 mmol kg^{-1} of soil) can be exceeded quite readily. However, Norvell (1984) proposed using a modified method with a 5:1 extractant to soil ratio for acid and metal–contaminated soils.

EDTA based extractions have been used for longer than the DTPA based extractions, with references dating back to the 1950s (Cheng and Bray 1953; Viro 1955a, b). Unlike the DTPA extraction of Lindsay and Norvell (1978) there is no single usual extraction with standardized concentration (e.g., in the range 0.1–0.01 M) and standardized pH value. There are many documented cases of EDTA extractions correlating well with plant metal concentrations (e.g., Cajuste and Laird 2000; Hooda et al. 1997; Michaud et al. 2007).

A rather different approach worthy of mention is *diffusive gradient in thin films* (DGT). The DGT methodology was originally developed for assessing water chemistry (Davison and Zhang 1994), but has more recently been applied to measure bioavailability in soils, e.g., Zhang et al. (2001). The technique involves applying an arrangement of a filter, layer of diffusive gel and layer of resin to a moist soil surface. Over time, contaminants diffuse through the gel and are adsorbed by the resin. This approach samples both contaminants in solution and those held loosely on exchange sites which come into solution to replace contaminants that are sorbed by the resin. Zhang et al. (2001) report good correlations between copper uptake by *Lepidium heterophyllum* and the concentration of Cu measured using DGT. Good correlations between plant and DGT concentrations have also been reported for copper in other studies (e.g. Song et al. 2004) and zinc (e.g. Cornu and Denaix 2006; Sonmez and Pierzynski 2005). Cornu and Denaix (2006), however, found a weak correlation between plant and DGT concentrations, for cadmium. The use of DGT for organisms is debated, for instance Koster et al. (2005), suggest that the DGT is often poorly representative for the metal concentrations to which the organism is exposed. In general, DGT will deplete the pore water of metals and replenishment of metals towards the solution phase is a slow process. However, organisms move around and are likely exposed to concentrations that are controlled by equilibrium between pore water and the solid phase, while metal concentrations are not depleted.

16.4.1.2 Organic Contaminants

Organic contaminants have been recognised as contaminants in soils for a shorter period of time than inorganic contaminants. This is due to the historic importance of metals, the relative ease of analysis of metals compared to organic contaminants, and the mere fact that the spectrum of chemical structures is more diverse for organic contaminants than for metals and metalloids. Although the list of potential approaches to simulate bioavailability of organic contaminants is long, consequently far fewer chemical extraction procedures have been proposed as proxies for the availability of organic contaminants compared to metals. Paradoxically, this makes extraction procedures for organic contaminants harder to summarise and review, because there has been less time for the establishment of one or several standardized procedures. Additionally, compared to metal extractions, a far higher proportion of extractions have been validated to organisms other than plants. A good review of different approaches is given by Ehlers and Loibner (2006) and readers are referred to that paper for details.

As with metals a number of physical-chemical (extraction) procedures have been developed to aid in the prediction of the bioavailable concentration of organic contaminants. These procedures provide knowledge about the extent of contaminant retention in soils, within shorter periods and for lower budgets and give more precise information on soil constituents being responsible for the sequestration of hydrophobic organic contaminants in soil compared to bioassays. Table 16.3 provides an outline of some of the principle studies reported in the literature in which chemical solvents were used to evaluate bioavailability/bioaccessibility of organic contaminants. The most straightforward extractions involve shaking soil with a weakly polar (e.g., methanol-water, *n* butanol, ethanol) or non-polar (e.g., hexane) liquid. Kelsey and Alexander (1997) and Tang and Alexander (1999) report good correlations between extractable concentrations of atrazine, phenanthrene, anthracene, fluoranthene and pyrene and uptake by earthworms and plants and breakdown by bacteria, but found no universal extractant that produced good correlations for all chemicals and organisms tested.

In recent years, biomimetic extractions have become increasingly common to assess bioavailability of organic contaminants in soils. Among biomimetic extractions of hydrophobic organic contaminants, two distinctive lines of approaches can be distinguished, i.e. those that are equilibrium-based and negligibly change the partitioning of a contaminant between the solid phase and the pore water, and those that deplete all contaminants released from the solid phase into the pore water, within a certain time period.

Table 16.3 Outline of some principal studies that employed chemical solvents to evaluate bioavailability of organic contaminants

Contaminant	Solvent	Bioassay	Operation	Comments
Atrazine, Phenanthrene (Kelsey and Alexander 1997)	Methanol/water, n-Butanol	Earthworm uptake and degradation.	25 ml solvent and 10 g solid. Shaking for 2 h.	Methanol/water best predictor for atrazine whereas n-butanol was best for phenanthrene.
DDT, DDE, DDD PAH (mixture) (Tang et al. (1999, 2002))	THF, Ethanol	Earthworm uptake.	15–20 ml solvent and 1 g soil, 10 s of mixing.	Good correlation with earthworm accumulation.
Anthracene, Fluoranthene, Pyrene (Tang and Alexander (1999))	n-Butanol Propanol Ethyl acetate	Plant retention, earthworm uptake and microbial degradation.	25 ml solvent and 1–2 g soil, 5 s of mixing.	Reasonable correlation with bioassays.
Phenanthrene, Pyrene, Chrysene (Liste and Alexander (2002))	n-Butanol	Earthworm uptake and microbial degradation.	15 ml extractant, 5–10 g soil, mixing: 5 s (worm) or 120 s (degradation).	Applicable for bioavailability prediction.

Taken from Jensen and Mesman (2006)

Non-depleting procedures measure chemical activity of a contaminant in soil, whereas depleting procedures measure the bioaccessible fraction of a contaminant. Examples of equilibrium based procedures are negligibly depleting solid phase micro extraction (nd-SPME) (see details later) and extraction with polyoxymethylene strips, triolein-embedded cellulose acetate membranes, and hollow fibre supported liquid membranes. Depleting extraction procedures include, for example, Tenax beads and β-cyclodextrin. The later biomimetic extractions are based on the principle that organisms predominantly take up contaminants from the pore water and thereby temporarily deplete the pore water, which again is rapidly replenished with contaminants desorbing from the solid phase.

Reid et al. (2000) propose shaking soil with an aqueous solution of hydroxypropyl-β-cyclodextrin (HPCD), a large molecule with a hydrophilic exterior surface and a hydrophobic cavity within the molecule. Reid et al. (2000) report good correlation between HPCD extractable polycyclic aromatic hydrocarbons (PAHs) and biodegradation of PAHs (and thus bioavailability) by microorganisms. Cuypers et al. (2002) also report good correlations. An alternative approach to assess bioavailability/bioaccessibility has been the use of C18 membrane disks, which are placed in the soil and then removed and purged. The amounts of contaminant extracted have shown good correlations between, for example, extractable DDT, DDE, DDD and earthworm uptake (Tang et al. 1999), extractable PAH and earthworm uptake (Tang et al. 2002) and extractable PAH and PCB and earthworm uptake (Krauss and Wilke 2001). Beads made from Tenax TA have also been suggested as an extractant (Cornelissen et al. 1998). A slurry of soil, Tenax TA beads and salt solution is shaken for a specified period of time and then the Tenax TA beads (which float or stick to the vessel walls) are removed and extracted for organic contaminants. Morrison et al. (2000) report good correlations between Tenax TA extracted DDT, DDE and DDD and earthworm tissue concentrations, but poor correlations for dieldrin. Cuypers et al. (2002) report good correlations between Tenax extractable PAH and biodegradation of PAHs. Ten Hulscher et al. (2003) also report good correlations between extractable PAHs and uptake in earthworms, though the nature of the relationship varied with exposure media. Results reported by De La Cal et al. (2008) suggest that 48 h of Tenax extraction is needed as proxy for the bioaccessible fraction of highly hydrophobic organic contaminants like polybrominated diphenyl ethers, DDT, and DDT metabolites. Finally, a proposed procedure for measuring bioavailability using a solid phase extractant is “Solid-phase micro-extraction with negligible depletion” (nd-SPME) which involves inserting a fibre thinly coated in an organic compound such as poly(dimethysiloxane) (PDMS) and polyacrylate into the soil, leaving it there for a period of time and then removing it and extracting organic contaminants from the fibre. Van der Wal et al. (2004) report good correlations between accumulation of HCB, telodrin, dieldrin and PCBs in earthworms and extractions using SPME.

16.4.2 Modelling the Bioavailability of Contaminants

16.4.2.1 Metals and Metaloids

Models can be split into two varieties, mechanistic models (i.e. those with a theoretical basis) and empirical models (i.e. those which are correlations). An example of a model based on theory is the *biotic ligand model*.

The biotic ligand model is used to predict the toxicity of contaminants to target organisms. It was initially designed for aquatic systems (e.g., Di Toro et al. 2001), but more recently has also been developed for soil systems (e.g., Allen et al. 2008; Lock et al. 2007; Steenbergen et al. 2005; Thakali et al. 2006a, 2006b; Van Gestel and Koolhaas 2004). The majority of chemical extractions that are used as proxies for bioavailability are correlated with the metal concentration in an organisms tissue. In contrast, the biotic ligand model is used to predict a toxicological endpoint, be that root elongation, earthworm or springtail reproduction, respiration et cetera. Thus biotic ligand models have great potential for the assessment of ecological risks related to contaminated sites, as they determine not just whether a contaminant is bioavailable, but also whether that contaminant will have a toxic effect. Although from a contaminated site Risk Assessment perspective the toxicity component of biotic ligand models is their most valuable component and it is almost impossible to separate the bioavailability and toxicity components, the theory behind the biotic ligand model is important as an approach to modelling bioavailability and is described below.

The theory underlying the biotic ligand model is that free metal ions react with binding sites at the organism – pore water interface and the fraction of binding sites occupied by the metal of interest governs the toxic response. There is competition for these binding sites between the contaminant of interest and other ions present such as H^+, Mg^{2+} and Ca^{2+}. Thus the bioavailability component of the model comprises determining the free metal ion activity in the pore water and determining the partition coefficient between the pore water and the binding sites of the organism (the biotic ligand). Free metal ion activities are either measured directly with an ion specific electrode (e.g. Steenbergen et al. 2005), or calculated using a chemical speciation program such as WHAM (Windemere Humic Acid Model) VI (e.g. Thakali et al. 2006a, b; Tipping 1998). The partition coefficients are derived by experiments in which toxic effects and ion activities are measured with one variable, e.g., activity of the metal of interest or pH being varied whilst other variables are kept constant (see De Schamphelaere and Janssen 2002). It is important to note that the above approach is different from simply correlating metal free ion activity with the metal concentration in the organism of interest. Sometimes such correlations exist, but not always as competition exists between the metal of interest and other ions for the exchange sites on the biotic ligand.

Empirical models generally measure the tissue concentration of a metal in the organism of interest and a host of soil properties, such as pH, bulk metal concentration, pore water metal concentration, dissolved organic carbon, concentration of Fe and Al oxyhydroxides, concentration of clay minerals, et cetera. Multiple linear regression techniques are then used to derive a predictive relationship for tissue concentrations. A large number of such models exist in the literature for earthworms and were recently reviewed by Nahmani et al. (2007b). The majority of these models take the form:

$$\mathrm{Log}\, M_{ew} \approx \mathrm{a} * \mathrm{Log}\, M_s + \mathrm{b} \tag{16.6}$$

where:

M_{ew} = concentration of metal in the earthworm (mg kg^{-1})
M_s = concentration of metal in the soil (mg kg^{-1})

The problem with these models (be they for earthworms or other organisms) is that they are hardly ever, if at all, validated with independent data sets, largely due to the lack of appropriate data. Exposure periods, for example, may differ between experiments, different species of test organisms may have been used, a variable present in the regression may not have been measured in another study, et cetera. Thus, their applicability to different sites and soils is always open to question. One of the few studies in which relationships were derived and validated with independent data is that of Sample et al. (1999). In this study 26 data sets were used to derive regression equations and six data sets to validate the data. The best regression coefficients were obtained when metal body burden was regressed against soil concentration and (usually) soil Ca-concentration. Including pH in the regressions occasionally improved the predictive power of the equations. Nahmani et al. (2009) applied a selection of these equations and those derived by Neuhauser et al. (1995) and Peijnenburg et al. (1999) to their own data set that they had used to determine rates of metal uptake. The equations of Sample et al. (1999) were found to best predict tissue metal concentrations.

16.4.2.2 Organic Contaminants

As with metals, for organic contaminants a distinction can be made between statistical models linking accumulated contaminants to extractable concentrations or pore water contents, and more mechanistically based modelling approaches like the one reported by Jager (1998) for estimating bioconcentration in earthworms.

Three conceptual frameworks provide the basic concepts for modelling bioavailability of contaminants. The first conceptual framework is the concept of chemical equilibrium in which chemical activities (or fugacities) are the driving factor for transport and distribution processes, including passive uptake of contaminants by biota (Diamond et al. 1992). The fugacity concept dictates chemical fugacities to be similar across biological membranes and explains observed variability in uptake patterns for organisms for which active uptake (like feeding and ingestion of solid soil particles) is of importance. Similarly, the concept explains why deviations from pore water uptake are often observed for highly hydrophobic contaminants (i.e., with log-transformed values of the octanol-water partition coefficient > approximately 5). Within the fugacity concept, Reichenberg and Mayer (2006) identified two complementary aspects of bioavailability of organic contaminants, these being the accessible quantity and the chemical activity that is to be deduced from this quantity as related to the physico-chemical conditions of the soil.

The second conceptual framework is the concept of equilibrium partitioning theory (EPT) in which chemical activities in the pore water are assumed to drive uptake and effects (Van Gestel 1997). The equilibrium partitioning concept is schematically given in Fig. 16.2. EPT assumes that the major distribution processes in the soil compartment, i.e., between soil – pore water – biota, can be described or predicted from simple physico-chemical properties such as the lipophilicity of the organic contaminant and the relative amount of binding sites in the soil or fat content of the biota. The EPT can be used to predict body residues and/or toxicity in soil dwelling species on the basis of data generated with aquatic organisms assuming that either pore water is the only route of exposure or that additional uptake pathways are proportional to pore water uptake, and that aquatic species generally have the same overall sensitivity distribution as terrestrial species. Within the basic EPT concept it is explicitly realized that the morphology, physiology and behaviour of biota dominate actual uptake and effects.

The third and most general concept is the general concept of bioavailability advocated by the ISO-working group on bioavailability (Harmsen 2007; ISO/DIS 2006). This group established a general bioavailability scheme for both organic and inorganic contaminants.

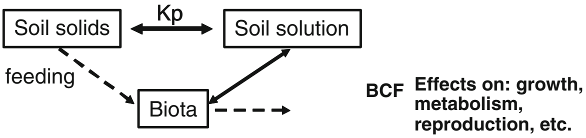


Fig. 16.2 Schematic representation of the equilibrium partitioning theory (EPT) (Kp = partitioning coefficient, BCF = bioconcentration factor)

It should be noted that the fugacity concept and the concept of equilibrium partitioning share many communalities and in fact, Reichenberg and Mayer (2006) added another concept: the thermodynamic concept of the chemical potential. In practical terms, the concepts have in common that it is the concentration of the contaminant in the pore water that is to be used as the basis for modelling uptake and effects of contaminants.

A model developed by Belfroid et al. (1995) in which estimates of accumulation of organic contaminants by earthworms (inputs: bioconcentration in water, sorption and elimination constants, rate of soil ingestion, and uptake efficiency) were generated, showed that in most cases uptake from pore water is the dominant uptake route. However, for contaminants with Log K_{ow} exceeding 5, oral uptake may contribute significantly: about 10% for soils with an average organic matter content of 3%, and about 50% for soils with an organic matter content of 20% (Belfroid et al. 1995). This implies that at maximum, the modelled concentration of organic contaminants exceeds the EPT prediction by a factor of 2. As reported by Jager (2003), this is similar to the differences found in general when comparing values for the bioaccumulation factor (BAF) and the bioconcentration factor (BCF). The BAF and BCF represent one of the most simplified single-compartment models for bioaccumulation, predicting partitioning between exposure medium and biota. Both BCF and BAF are generally calculated as the ratio, at equilibrium, of internal biota concentration to exposure concentration, with BCF being related to accumulation in organisms arising from water only, and BAF to accumulation from water and dietary sources (McGeer et al. 2003). In general, BAF is derived from measurements in the field and BCF is more readily measured under laboratory conditions.

Further support, although not for a soil inhabiting species, is obtained from experiments with *Lumbriculus variegatus*. Sormunen et al. (2008) showed that these sediment-ingesting worms have access to an additional bioavailable fraction of PCBs on top of the pore water pool that was especially evident when pore-water concentrations of PCB 77 approached the solubility limit. Thus, feeding may modify the bioavailable fraction in a way that cannot be explained by simple equilibrium partitioning models.

In laboratory tests, the validity of the EPT has been confirmed for various organisms like earthworms, enchytraeids, and nematodes. Almost the same evidence has been found for collembolans and isopods, while no prediction was possible for snails due to specific uptake mechanisms. Thus, evidence exists that pore water is the main uptake route for the above-mentioned organism taxa, at least for contaminants with a log $K_{ow} < 6$. For contaminants with a log $K_{ow} > 6$ other routes, e.g., via feeding, become more relevant for the same organisms.

16.5 Concluding Remarks

Bioavailability is an important concept when considering the ecological risks posed by contaminated sites. Legislators are moving in the direction of incorporating measures of bioavailability into regulations regarding contaminated sites. In particular concepts of bioavailability are important in higher tier Risk Assessments. Whilst bioavailability, and related terms such as bioaccessibility, can be precisely defined, the impact that soil properties have on soils and the multitude of organisms found at contaminated sites means that a “one size fits all” precise definition of bioavailability only exists on a philosophical level. Any given contaminant will have differing bioavailabilities to different organisms in different soils. With all this potential variability it is not surprising that a range of different methods for measuring and calculating bioavailability exist. The chemical methods which exist for predicting bioavailability are validated for specific conditions, but can not hope to be universally applicable. Similarly, models have been developed which work well for specific conditions, but are inevitably organism-specific. Thus when applying proxies for bioavailability, it is important to bear in mind the conditions for which that proxy was developed. At present it is the case that the best method to determine contaminant concentrations and effects in organisms is to measure them. Chemical extractions are probably currently the better option for predicting longer term bioavailability providing that the drawbacks of extractions are clearly stated. However, in the future calculations will replace chemical extractions as the method of choice for predicting bioavailability. In terms of using bioavailability measures to support contaminated site ecological Risk Assessment and Risk Management, legislators must identify key organisms, with the support of ecotoxicologists, that they wish to protect, for example earthworms as an example of a major soil invertebrate. Once key species have been identified there is cause to be optimistic that suitable methods for assessing bioavailability either exist or can be developed. A key need for legislators is not additional and new methods for measuring bioavailability, but a series of validation studies using existing methods.

References

Allen HE, Yanqing L, Di Toro DM (2008) Ecotoxicity of Ni in soil. Mineral Mag 72:367–371

Arnold RE, Hodson ME, Comber S (2007) Effect of organic complexation on the toxicity of Cu to the earthworm *Eisenia fetida*. Appl Geochemi 22:2397–2405

Belfroid A, Seinen W, Van Gestel CAM, Hermens JLM, Van Leeuwen K (1995) Modelling the accumulation of hydrophobic organic chemicals in earthworms – application of the equilibrium partitioning theory. Environ Sci Pollut Res 1:5–15

Binger CA, Martin JP, Allen-King RM, Fowler M (1999) Variability of chlorinated-solvent sorption associated with oxidative weathering of kerogen. J Contamin Hydrol 40:137–158

Bintein S, Devillers J (1994) Qsar for organic-chemical sorption in soils and sediments. Chemosphere 28:1171–1188

Brun LA, Maillet J, Hinsinger P, Pepin M (2001) Evaluation of copper availability to plants in copper-contaminated vineyard soils. Env Pollut 111:293–302

Bucheli T, Gustaffson O (2000) Quantification of the Soot-Water Distribution Coefficient of PAHs Provides Mechanistic Basis for Enhanced Sorption Observations. Environ Sci Technol 34:5144–5151

Cajuste LJ, Laird RJ (2000) The relationship between phytoavailability and the extractability of heavy metals in contaminated soils. In: Iskandar IK (ed) Environmental restoration of metals-contaminated soils. Lewis, Boca Raton, FL, pp 189–198

Campbell PGC, Chapman PM, Hale BA (2006). Risk assessment of metals in the environment. Issues in environmental science and technology, no. 22. Chemicals in the environment: assessing and managing risk © The Royal Society of Chemistry

Cheng KL, Bray RH (1953) Two specific methods of determining copper in soil and plant material. Anal Chem 25:655–659

Conder JM, Lanno RP, Basta NT (2001) Assessment of metal availability in smelter soil using earthworms and chemical extractions. J Environ Qual 30:1231–1237

Cornelissen G, Rigterink H, Ferdinandy MMA, Van Noort PCM (1998) A simple TENAX® extraction method to determine the availability of sediment-sorbed organic compounds. Environ Toxicol Chem 20:706–711

Cornu JY, Denaix L (2006) Prediction of zinc and cadmium phytoavailability within a contaminated agricultural site using DGT. Environ Chem 3:61–64

Cuypers C, Pancras T, Grotenhuis T, Rulkens W (2002) The estimation of PAH bioavailability in contaminated sediments using hydroxypropy-β-cyclodextrin and Triton X-100 extraction techniques. Chemosphere 46:1235–1245

Davison W, Zhang H (1994) In-situ speciation measurements of trace components in natural waters using thin-film gels. Nature 367:546–548

Dean JR (2007) Bioavailability, bioaccessibility and mobility of environmental contaminants. Wiley, Chichester, p 292

De La Cal A, Eljarrat E, Grotenhuis T, Barceló D (2008) Tenax® extraction as a tool to evaluate the availability of polybrominated diphenyl ethers, DDT, and DDT metabolites in sediments. Environ Toxicol Chem 27:1250–1256

De Schamphelaere KAC, Janssen CR (2002) A biotic ligand model predicting acute copper toxicity for *Daphnia magna*: the effects of calcium, magnesium, sodium, potassium and pH. Environ Sci Technol 63:48–54

Diamond ML, MacKay D, Welbourn PM (1992) Models of multimedia partitioning of multispecies chemicals – the fugacity equivalence approach. Chemosphere 25:1907–1921

Di Toro DM, Allen HE, Bergman HL, Meyer JS, Paquin PR, Santore RC (2001) Biotic ligand model of the acute toxicity of metals.1. Technical basis. Environ Toxicol Chem 20: 2383–2396

Ehlers GAC, Loibner AP (2006) Linking organic pollutant (bio)availability with geosorbent properties and biomimetic methodology: a review of geosorbent characterisation and (bio)availability prediction. Environ Pollut 141:494–512

Ehlers LJ, Luthy RG (2003) Contaminant bioavailability in soil and sediment. Environ Sci Technol 37:295A–302A

European Commission (2003) Technical guidance document on risk assessment, part III. European commission document reference EUR 20418/EN/3. Luxembourg

Fang J, Wen B, Shan XQ, Lin JM, Owens G (2007) Is an adjusted rhizosphere-based method valid for field assessment of metal phytoavailability? Application to non-contaminated soils. Environ Pollut 150:209–217

Faure G (1992) Principles and applications of inorganic geochemistry. Macmillan, New York

Grathwohl P (1990) Influence of organic-matter from soils and sediments from various origins on the sorption of some chlorinated aliphatic-hydrocarbons – implications on K_{OC} correlations. Environ Sci Technol 24:1687–1693

Haines BL, Waide JB, Todd RL (1982) Soil solution nutrient concentrations sampled with tension and zero-tension lysimeters: report of discrepancies. Soil Sci Soc Am J 46:547–555

Harmsen J (2007) Measuring bioavailability: from a scientific approach to standard methods. J Environ Qual 36:1420–1428

Hooda PS, McNulty D, Aalloway BJ, Aitken MN (1997) Plant availability of heavy metals in soils previously amended with heavy applications of sewage sludge. J Sci Food Agric 73 446–454

Houba VJG, Lexmond ThM, Novozamsky I, van der Lee JJ (1996) State of the art and future developments in soil analysis for bioavailability assessment. Sci Total Environ 178:21–28

ISO (2005) 11074: Soil quality – vocabulary. ISO, Geneva, Switzerland

ISO/DIS (2006) 17402: Soil quality. Guidance for the selection and application of methods for the assessment of bioavailability of contaminants in soil and soil materials. ISO, Geneva, Switzerland

Jager T (1998) Mechanistic approach for estimating bioconcentration of organic chemicals in earthworms (*oligochaeta*). Environ Toxicol Chem 17:2080–2090

Jager DT (2003) Worming your way into bioavailability – modelling the uptake of organic chemicals in earthworms. PhD Thesis, University of Utrecht, The Netherlands

Jensen J, Mesman M (eds) (2006) Ecological risk assessment of contaminated land – decision support for site specific investigations. RIVM report 711701047. Dutch National Institute of Public Health and the Environment, Bilthoven, The Netherlands. www.rivm.nl

Jones OAH, Voulvoulis N, Lester JN (2006) Partitioning behavior of five pharmaceutical compounds to activated sludge and river sediment. Arch Environ Contamin Toxicol 50:297–305

Karickhoff SW, Brown DS, Scott TA (1979) Sorption of hydrophobic pollutants on natural sediments. Water Res 13:241–248

Kelsey JW, Alexander M (1997) sepective chemical extract to predict bioavailability of soil-aged organic chemicals. Environ Sci Technol 31:214–217

Koster M, Reijnders L, Van Oost NR, Peijnenburg WJGM (2005) Comparison of the method of diffusive gels in thin films with conventional extraction techniques for evaluating zinc accumulation in plants and isopods. Environ Poll 133:103–116

Krauss M, Wilke W (2001) Biomimetic extraction of PAHs and PCBs from soil with octadecyl-modified silica disks to predict their availability to earthworms. Environ Sci Technol 35: 3931–3935

Krogh L (2000) Cation-exchange capacity pedotransfer functions for Danish soils. Acta Agric Scand B 50:1–12

Langdon CJ, Piearce TG, Meharg AA, Semple KT (2001) Resistance to copper toxicity in populations of the earthworms *Lumbricus rubellus* and *Dendrodrilus rubidus* from contaminated mine wastes. Environ Toxicol Chem 20:2336–2341

Lindsay WL (1979) Chemical equilibria in soils. Wiley, New York

Lindsay WL, Norvell WA (1978) Development of a DTPA soil test for zinc, iron, manganese and copper. Soil Sci Soc Am J 42:421–428

Liste HH, Alexander M (2002) Butanol extraction to predict bioavailability of PAHs in soil. Chemosphere 46:1011–1017

Lock K, Van Eeckhout H, De Schamphelaere KAC, Criel P, Janssen CR (2007) Development of a biotic ligand model (BLM) predicting nickel toxicity to barley (*Hordeum vulgare*). Chemosphere 66:1346–1352

MacDonald JD, Bélanger N, Sauvé S, Courchesne, F, Hendershot WH (2008) Collection and characterization of soil solutions. In: Carter MR, Gregorich EG (eds) Soil sampling and methods of analysis. Canadian society of soil science. CRC, Boca Raton, FL, pp 179–196

McBride MB (1994) Environmental chemistry of soils. Oxford University, Oxford

McGeer JC, Brix KV, Skeaff JM, Deforest DK, Brigham SI, Adams WJ, Green A (2003) Inverse relationship between bioconcentration factor and exposure concentration for metals: implications for hazard assessment of metals in the aquatic environment. Environ Toxicol Chem 22:1017–1037

McLaughlin MJ, Zarcinas BA, Stevens DP, Cook N (2000) Soil testing for heavy metals. Commun Soil Sci Plant Anal 31(11–14):1661–1700

Mench M, Baize D, Mocquot B (1997) Cadmium availability to wheat in five soil series from the Yonne district, Burgundy, France. Environ Pollut 95:93–103

Michaud AM, Bravin MN, Galleguillos M, Hinsinger P (2007) Copper uptake and phytotoxicity as assessed in situ for durum wheat (*Triticum turgidum durum* L.) cultivated in Cu-contaminated, former vineyard soils. Plant Soil 298:99–111

Morrison DE, Robertson BK, Alexander M (2000) Bioavailability to earthworms of aged DDT, DDE, DDD and Dieldrin in soil. Environ Sci Technol 34:709–713

Nahmani JY, Hodson ME, Black S (2007a) Effects of metals on life cycle parameters of the earthworm *Eisenia fetida* exposed to field-contaminated, metal-polluted soils. Environ Pollut 149:44–58

Nahmani JY, Hodson ME, Black S (2007b) A review of studies performed to assess metal uptake by earthworms. Environ Pollut 145:402–424

Nahmani JY, Hodson ME, Devin S, Vijver M (2009) Uptake and excretion kinetics of metals by the earthworm *Eisenia fetida* exposed to field contaminated soils. Environ Pollut 157:2622–2628

National Research Council (2002) Bioavailability of contaminants in soils and sediments: processes, tools and applications. National Academies, Washington, DC

Neuhauser EF, Cukic ZV, Malecki MR, Loehr RC, Durkin PR (1995) Bioconcentration and biokinetics of heavy-metals in the earthworm. Environ Pollut 89:293–301

Norvell WA (1984) Comparison of chelating agents as extractants for metals in diverse soil materials. Soil Sci Soc Am J 48:1285–1292

O'Connor GA (1988) Use and misuse of the DTPA soil test. J Environ Qual 17:715–718

Peijnenburg WJGM, Baerselman R, de Groot AC, Jager T, Posthuma L, Van Veen RPM (1999) Relating environmental availability to bioavailability: soil type dependent metal accumulation in the oligochaete *Eisenia andrei.* Ecotoxicol Environ Safety 44:294–310

Perez-de-Mora A, Madejon E, Burgos P, Cabrera F (2006) Trace element availability and plant growth in a mine spill-contaminated soil under assisted natural remediation II. Plants. Sci Total Environ 363:38–45

Reichenberg F, Mayer P (2006) Two complementary sides of bioavailability: accessibility and chemical activity of organic contaminants in sediments and soils. Environ Toxicol Chem 25:1239–1245

Reid BJ, Stokes JD, Jones KC, Semple KT (2000) Nonexhaustive cyclodextrin-based extraction technique for the evaluation of PAH bioavailability. Environ Sci Technol 34:3174–3179

Ross SM (1994) Retention, transformation and mobility of toxic metals in soils. In: Ross SM (ed) Toxic metals in soil-plant systems. Wiley, Chichester, pp 63–152

Sabljić A, Güsten H, Verhaar H, Hermens J (1995) QSAR modelling of soil sorption. Improvements and systematics of log K_{OC} vs. log K_{OW} correlations. Chemosphere 31:4489–4514

Sample BE, Suter GW, Beauchamp JJ, Efroymson RA (1999) Literature-derived bioaccumulation models for earthworms: development and validation. Environ Toxicol Chem 18:2110–2120

Schröder TJ, Hiemstra T, Vink JPM (2005) Modeling of the solid-solution partitioning of heavy metals and arsenic in embanked flood plain soils of the rivers Rhine and Meuse. Environ Sci Technol 39:7176–7184

Semple KT, Doick KJ, Jones KC, Burauel P, Craven A, Harms H (2004) Defining bioavailability and bioaccessibility of contaminated soil and sediment is complicated. Environ Sci Technol 38:228A–231A

Song J, Zhao FJ, Luo YM, McGrath SP, Zhang H (2004) Copper uptake by Elsholtzia splendens and Silene vulgaris and assessment of copper phytoavailability in contaminated soils. Environ Pollut 128:307–315

Sonmez O, Pierzynski GM (2005) Assessment of zinc phytoavailability by diffusive gradients in thin films. Environ Toxicol Chem 24:934–941

Sormunen AJ, Leppanen MT, Kukkonen JV (2008) Influence of sediment ingestion and exposure concentration on the bioavailable fraction of sediment-associated tetrachlorobiphenyl in oligochaetes. Environ Toxicol Chem 27:854

Steenbergen NTTM, Iaccino F, de Winkel M, Reijnders L, Peijnenburg WJGM (2005) Development of a biotic ligan model and a regression model predicting acute copper toxicity to the earthworm *Aporrectodea caliginosa*. Environ Sci Technol 39:5694–5703

Stevenson FJ (1994) Humus chemistry: genesis, composition and reactions. Wiley, New York

Tang J, Alexander M (1999) Mild extractability and bioavailability of polycyclic aromatic hydrocarbons in soil. Environ Toxicol Chem 18:2711–2714

Tang J, Liste H-H, Alexander M (2002) Chemical assays of availability to earthworms of polycyclic aromatic hydrocarbons in soil. Chemosphere 48:35–42

Tang J, Robertson KB, Alexander M (1999) Chemical-extraction methods to estimate bioavailability of DDT, DDE and DDD in soil. Environ Sci Technol 33:4346–4351

Ten Hulscher TEM, Postma J, den Besten PJ, Stroomberg GJ, Belfroid A, Wegener JW, Faber JH, van der Pol JJC, Hendriks AJ, van Noort CM (2003) Tenax extraction mimics benthic and terrestrial bioavailability of organic compounds. Environ Toxicol Chem 22: 2258–2265

Thakali S, Allen HE. Di Toro DM, Ponizovsky AA, Rooney CP, Zhao F-J, McGrath SP (2006a) A terrestrial biotic ligand model. 1. Development and application to Cu and Ni toxicities to barley root elongation in soils. Environ Sci Technol 40:7085–7093

Thakali S, Allen HE. Di Toro DM, Ponizovsky AA, Rooney CP, Zhao F-J, McGrath SP, Criel P, van Eeckhout H, Janssen C, Oorts K, Smolders E (2006b) A terrestrial biotic ligand model. 2. Application of Ni and Cu toxicities to plants, invertebrates and microbes in soil. Environ Sci Technol 40:7094–7100

Tipping E (1998) Humic ion-binding model VI: an improved description of the interactions of protons and metal ions with humic substances. Aquatic Geochem 4:3–48

Tipping E, Rieuwerts J, Pan G, Ashmore MR, Lofts S, Hill MTR, Farago ME, Thornton I (2003) The solid-solution partitioning of heavy metals (Cu, Zn, Cd, Pb) in upland soils of England and Wales. Environ Pollut 125:213–225

Van der Wal L, Jager T, Fleuren RHLJ, Barendregt A, Sinnige TL, van Gestel CAM, Hermens JLM (2004) Solid-phase microextraction to predict bioavailability and accumulation of organic micropollutants in terrestrial organisms after exposure to a field-contaminated soil. Environ Sci Technol 38:4842–4848

Van Gestel CAM (1997) Scientific basis for extrapolating results from soil ecotoxicity tests to field conditions and the use of bioassays. In: Van Straalen NM, Løkke H (eds) Ecological risk assessment of contaminants in soil. Chapmann & Hall, London, pp 25–50

Van Gestel CAM, Koolhaas JE (2004) Water-extractability, free ion activity, and pH explain cadmium sorption and toxicity to *Folsomia candida* (Collembola) in seven soil-pH combinations. Environ Toxicol Chem 23:1822–1833

Vasudevan D, Cooper EM, Van Exem OL (2002) Sorption-Desorption of Ionogenic Compounds at the Mineral-Water Interface: study of Metal Oxide-Rich Soils and Pure-Phase Minerals. Environ Sci Technol 36:501–511

Viro PJ (1955a) Use of ethylenediaminetetraacetic acid in soil analysis I. Experimental Soil Sci 79:459–465

Von Oepen B, Kördel W, Klein W (1991) Sorption of nonpolar and polar compounds to soils: processes, measurements and experience with the applicability of the modified OECD-Guideline 106. Chemosphere 22:285–304

Viro PJ (1955b) Use of ethylenediaminetetraacetic acid in soil analysis II. Determination of soil fertility. Soil Sci 80:69–74

Williamson KL, Minard R, Masters KM (2007) Macroscale and microscale organic experiments, 5th edn. Houghton Mifflin, Chicago, IL

Zhang H, Zhao FJ, Sun B, Davison W, McGrath SP (2001) A new method to measure effective soil solution concentration predicts copper availability to plants. Environ Sci Technol 35: 2602–2607

Part V
Groundwater-Related Aspects

Chapter 17
Groundwater-Related Risk Assessment

Frank A. Swartjes and Juan Grima

Abstract Groundwater includes the pore water in the water-unsaturated upper soil layer as well as water usually referred to as groundwater, that is, the water in the water-saturated zone. Groundwater contains contaminants of natural and anthropogenic origin. It is generally recognised by all segments of society that fresh water is an immensely important resource. From a Risk Assessment point of view, groundwater needs to be approached from two different perspectives, namely, as an important protection target and as a means of transport (a pathway) for contaminants. For human beings the primary use for groundwater is as a source of drinking water. Although often underestimated, the water-saturated deeper soil layer is also a habitat for many organisms. For decades, there has been an on-going and interesting discussion concerning the intrinsic value of groundwater, sometimes including spiritual and even supernatural or religious arguments. Generally speaking, the transport of water and contaminants is much faster in the groundwater zone than in the water-unsaturated upper soil layer. Specific attention will be given in this chapter to the impact that a revised quantitative groundwater regime, the presence of heterogeneous soils or aquifers, surface water bodies, anthropogenic subsurface processes and structures, and heterogeneous soils and aquifers all have on groundwater quality. Additional attention will be paid to sustainable protection of groundwater resources, Conceptual Models, mathematical (numerical) models, Risk Management (including Natural Attenuation and regional approaches), sampling and monitoring, lysimeters and column experiments, the impact of climate change, mingling groundwater plumes, risk perception and communication, and the European Water Framework Directive.

F.A. Swartjes (✉)
National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands
e-mail: frank.swartjes@rivm.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_17, © Springer Science+Business Media B.V. 2011

Contents

17.1 Introduction

17.1.1 Subsurface Water

Groundwater is water beneath the earth’s surface. It encompasses the largest amount of fresh water in the world, in fact a much greater portion than is stored in surface waters, that is, lakes and rivers combined. This groundwater is found everywhere on earth, in some regions close to the surface such as in wetlands or marsh areas, while in arid areas up to several hundreds of meters below ground. In many regions, groundwater is directly linked to surface waters, forming complete water systems. In highly permeable limestone catchments, rivers can be entirely supported by the groundwater discharge. Fossil groundwater bodies are not, or are rarely connected to surface water systems.

It is generally recognized by all segments of society that fresh water, although not always associated with groundwater, is an immensely important resource. Although 70% of the earth's surface is covered by water, most of it is the salt water of seas and oceans. Only 2.5% of all water on earth is fresh water, most of which exists as glaciers and permanent snowfields, leaving only 0.26% of all fresh water available to support mankind (Eurogeosurveys 2009). More than 90% of this small volume is in the form of groundwater. The need of water for a rapidly growing human population worldwide and for its activities such as non-controlled water use, pose serious threats, along with industrial and agricultural contamination, to sustaining this important volume of fresh water. At present, humans use some 10% of the renewable fresh water resource, but it can be very variable depending on time of year, and on geographical location (Optimum Population Trust 2009). Moreover, humans must share it with all the other species on earth.

Water is part of a spectacular cycle known as the *hydrological cycle* (aka: water cycle) that includes an exchange of water between the soil-groundwater system, surface waters and the atmosphere (see Fig. 17.1).

Of all water volumes that play a role in the hydrological cycle, water in the soil has the highest residence time. Water can reside in aquifers for hundreds of years; see Fig 17.2, which shows recharge and discharge areas, flow lines and residence time of water in an aquifer (López-Geta et al. 2006).

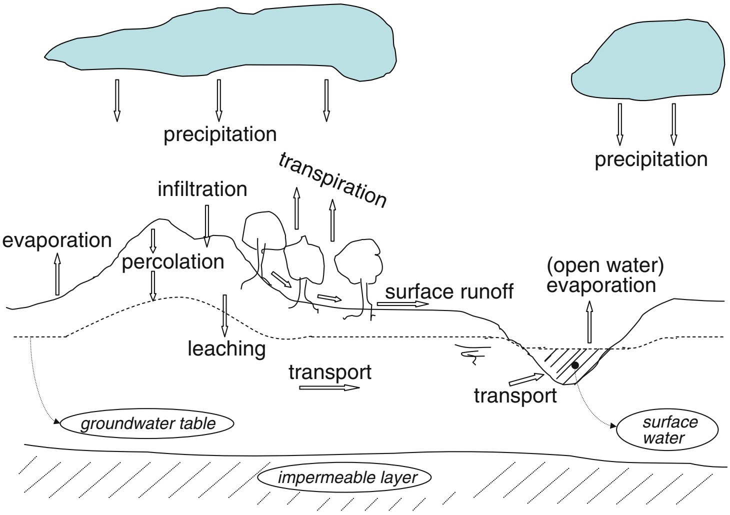


Fig. 17.1 The hydrological cycle, including an exchange of water between the soil-groundwater system, surface waters and the atmosphere

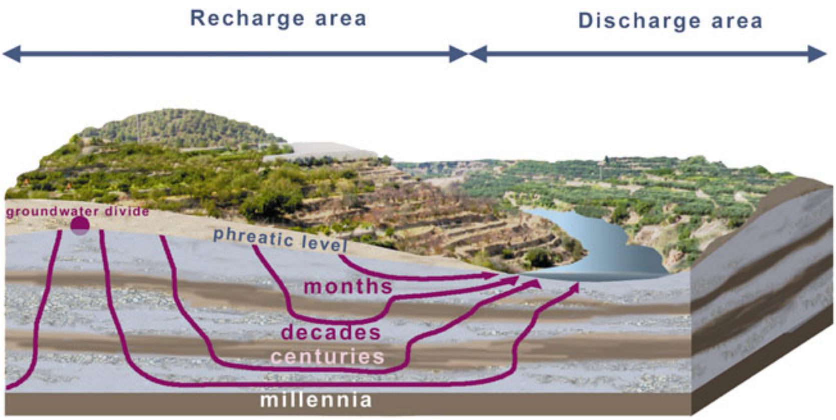


Fig. 17.2 Recharge and discharge areas, flow lines and residence time of water in an aquifer, from López-Geta et al. (2006; reproduced with permission)

It is expected that climate change will impact the hydrological cycle in the future, both in terms of magnitude and in regularity of fluxes. Several scenarios have been developed to predict the potential impact of climate change on the hydrological regime. Although there is no general consensus on the quantitative effects, it is widely accepted that many (semi-)arid areas will become drier, resulting in less groundwater recharge. In humid zones, recharge might increase due to a higher number of extreme precipitation events. Efforts are being undertaken so as to better understand the climate change mechanisms on a regional and even local scale, and hence to be able to determine mitigating measures.

In contrast to soil, groundwater is highly mobile. As a consequence, groundwater can transport contaminants from one place to another. Contaminants are transported from the upper soil into layers lying beneath it and, ultimately, into the groundwater zone and surface waters. Another implication of this mobility is that water, including groundwater, crosses borders, including international ones. This has legal consequences when contaminants migrate from the contaminated site, at which the contamination source is or was active, to another, clean site, often with a different jurisdiction and owner. Policy on groundwater should therefore have an interregional and an international dimension. An interesting phenomenon from this standpoint was the cooperation between the water authorities of the former city of West-Berlin with those of the former German Socialistic Republic, during the period of the division of West and East Germany before 1989. While the authorities of these completely different political systems pursued an almost totally separate policy in all other respects, the natural course of water (and, hence, of contaminant) migration forced the authorities to collaborate in the field of water management.

Since groundwater is used for industrial manufacture and in producing food, there is much ‘embedded’ groundwater exported around the world. With increasing shortages of water resources, this is a growing concern.

With regard to Risk Assessment, the conclusion that can be derived from what has been outlined above is that groundwater needs to be approached from two different perspectives. Firstly, groundwater is an important *protection target* as it may be threatened by contamination and, therefore, needs to be protected. Secondly, groundwater is an important *pathway*, that is, a means of transport (carrier) for contaminants.

17.1.2 Terminology

As described in the Introduction chapter (see Section 1.2.1), soil can be divided into two different entities, that is, a *water-unsaturated upper soil layer* (upper soil) and a *water-saturated groundwater zone*. These two entities are separated by a groundwater table. Most definitions describe groundwater as all water present beneath the soil's surface. In fact, according to this definition three types of groundwater can be distinguished, covering both entities described above. The first is the *pore water* in the water-unsaturated upper soil layer. This water shares the soil pores with soil gas and is partially unavailable, since it is attached to soil particles. The water volume (and, hence, gas volume) in the unsaturated zone changes according to relatively short time frames due to rain and snow events, evapotranspiraton and water flow. The second is the water usually referred to as *groundwater*, that is, the water in the water-saturated zone. These water volumes are extremely important, since they constitute the major source for drinking water in many countries. The major groundwater reservoirs are present in *aquifers*, that is, in layers that carry huge volumes of renewable or fossil water. In phreatic aquifers, a free water table exists at atmospheric pressure. In confined and artesian aquifers, the water is held in aquitards, that is, layers with poor water permeability. Therefore, the water pressure can rise above the top of the aquifer, even higher than the soil surface, creating free-flowing wells, under so-called artesian conditions. The third origin of groundwater is in so-called secondary porosity and permeability as found, for example, in fractures or karstic phenomenon in consolidated rocks, often at great depths beneath the soil surface. Although this groundwater occurrence is not often included in Risk Assessment, it can be an important resource at a local or even regional scale that needs to be protected, and can also play an often underestimated role as a pathway for contaminants (B. Harris 2009 'personal communication'). The reason for this is that, in regions with karstic or pseudo-karstic features, contamination can spread at much higher rates than would normally be expected for groundwater.

Remarkably, in the great majority of studies that focus on Risk Assessment related to contaminated aquifers, it is referred to as 'contaminated groundwater', while actually groundwater is only one of the two constituents of an aquifer. It often seems that the solid phase of the water-saturated soil has been forgotten in Risk Assessments. There are several reasons, however, to put this much emphasis on the groundwater compartment. Firstly, analogous to the solid phase compartment in the upper soil layer (for example, soil used as building material), groundwater has been directly used as a vital human resource since early mankind. Secondly, the solid phase of aquifers generally does not have as much capacity to adsorb contaminants as the organic-rich upper soil layers do. As a consequence, the majority of the contaminants in the water-saturated groundwater zone usually resides in the water phase, that is, in the groundwater. For the same reasons, and to standardise the terminology, the term 'groundwater' is often used throughout this book to mean the whole water-saturated soil layer.

There is also a practical reason for focusing on the groundwater and not on the solid phase of the water-saturated soil. Sampling of groundwater is relatively easy, since groundwater can be sucked up via water-sampling tubes. Taking aquifer samples, that is, soil samples in the deeper soil layers that include solid phase material and the groundwater, is a more difficult, time-consuming and costly activity, especially in the case of deep groundwater tables. For this reason, groundwater samples are taken in routine investigations and aquifer samples are only taken in exceptional circumstances.

17.1.3 Groundwater Quality

17.1.3.1 Natural Impact on Groundwater

Contaminants coming from various natural and anthropogenic sources may affect groundwater bodies. Therefore, in all countries of the world, contaminants are found in the groundwater. In fact, pure water only exists in the laboratory in the form of demineralised water, especially as distilled water.

Water travels much more slowly in aquifers than in surface water systems, and the time for infiltrating water and contaminants to migrate from the land surface to the point of discharge is measured in years, decades or centuries, depending on the type of contaminant, the aquifer characteristics and the length of the flow path. As a consequence, there is a great deal of time for intimate contact between the water and soil and rock material, which results in changes of water quality with regard to mineral composition. As a result, metals are commonly found in groundwater in practically every country in the world (e.g., US Environmental Protection Agency 2008). Given the wide variety in the geometry of hydrological systems, in soil and rock materials, and, hence, travel times, the naturally impacted groundwater quality (*natural background concentration*) is highly variable.

Naturally impacted groundwater may sometimes result in unacceptable risks to human health or the ecosystem. An example of this is the arsenic poisoning of large numbers of people drinking shallow groundwater in Bangladesh and West Bengal, India, which arose from a natural phenomenon, not industrial contamination (Rahman et al. 2001).

17.1.3.2 Anthropogenic Impact on Groundwater

In addition to contaminants from natural sources, a wide variety of the contaminants that originate from point sources, that is, small-scale sources that often originated in leakage or dumping, are frequently found in the more densely populated areas of the world. Sewage lagoons, landfills, leaking sewage systems (Lerner 2002), along with waste disposal facilities, for example, can negatively affect groundwater quality. Individual sewage treatment systems (septic systems) for homes in rural areas and buildings not connected to the city sewer system – although in a quantitative sense generally of limited importance – are another factor to consider when evaluating risks to groundwater resources. An example of the large scale anthropogenic impact on groundwater is described in Santos et al. (2002), who measured zinc, cadmium, lead and copper in Andalusian groundwater, in the south of Spain resulting from the Aznalcóllar accident. In April 1998, the failure of the tailings dam of the lead zinc mine at Aznalcóllar near Seville, Spain, released nearly five million cubic meters of toxic tailing slurries into the valley of the Guadiamar River. As a result, the alluvial aquifer of Rio Agrio became contaminated.

Moreover, organic contaminants are frequently found in groundwater in all developed countries worldwide. In fact, in all situations where chemicals are handled, the groundwater beneath the location will be contaminated to a greater or a lesser extent (Lerner and Tellam 1992). The most frequently found contaminants in the groundwater are mobile contaminants that are not readily attenuated in the groundwater. Amongst these are the groups of chlorinated hydrocarbons and petroleum hydrocarbons, denser (Dense Non-aqueous Phase Liquids; DNAPL) and less dense (Light Non-aqueous Phase Liquids; LNAPL) than water, respectively. Moran et al. (2007), for example, registered at least one VOC in 12% of the samples from 1,926 rural private wells in the USA. The seven most frequently detected VOCs were: trichloromethane, methyl *tert*-butyl ether, tetrachloroethene, dichlorodifluoromethane, methylbenzene, 1,1,1-trichloroethane, and 1,2-dibromo-3-chloropropane. Solvents and trihalomethanes were the most frequently detected VOC groups in private wells. The concentrations of the most detected VOCs were relatively small and only 1.4 percent of the samples had one or more VOC concentrations that exceeded a federally established drinking water standard or health criterion. Kelsh et al. (2003), in yet another example, registered perchlorate in Californian groundwater. During the last few years, MTBE (methyl *tert*-butyl ether) emerged as a frequently found contaminant in groundwater in the USA (e.g., Deeb et al. 2003) and in Europe. Kolb and Püttmann (2006), for example, showed detection frequencies of MTBE in groundwater in Germany that were 58% at 29 known fuel-contaminated sites, 63% at 67 urban sites and even for non-urban sites they were 24%.

With regard to immobile contaminants in the upper soil layers, leaching to groundwater may continue for many decades. Mandocdoc and Primo David (2008), for example, detected dieldrin in concentrations that exceeded the drinking water standards in several wells at the former United States military facility Clark Air Base in the Philippines, 16 years after dieldrin was last used.

In addition, diffuse sources contribute to contaminated groundwater. The use of fertilisers in agriculture, for example, is the main cause of high nitrate concentrations in groundwater (e.g., Tomer and Burkart (2003), who demonstrated the impact of different agricultural management procedures on nitrate leaching to groundwater in Iowa, USA).

17.1.3.3 Impact of a Revised Quantitative Groundwater Regime

Any changes in the quantitative groundwater regime, it should be realised, are also going to impact groundwater quality. In wetlands, since strongly dependent on groundwater, the increase in temperature, sea-level rise and changes in precipitation will strongly affect the hydrological regime. In many wetlands the groundwater table is very shallow and reduction of the groundwater level could lead to drying out of the wetland. In addition, groundwater-level lowering processes very often result in a general worsening of groundwater quality. Sea-level rise, on the other hand, could produce coastal erosion and sea water intrusion into coastal wetlands and deltas.

Similar effects may be seen in urban areas, where the water regime is influenced by the large contribution of covered areas, including built-up areas. However, Lerner (2002) showed that total urban recharge is similar to or higher than rural recharge for two main reasons: excess rainfall is often routed to groundwater through soakaways, and covered surfaces are frequently more permeable than they appear and so allow some infiltration while suppressing plant growth and evapotranspiraton.

The groundwater quantity and quality may also become modified as an indirect result of water losses in urban water supply and sewerage systems. Losses from urban water supply distribution networks result in an increase in groundwater recharge. In the Alicante village of Vergel (Spain), for example, more than 45% of the water circulating through the urban supply system was lost due to poor maintenance of the infrastructure (De la Orden Gómez 2006). And losses from sewage pipelines can lead to groundwater contamination. The leakage rates from water mains are typically 25% (De la Orden Gómez 2006). Both exfiltration from and infiltration into sewers occur with only a small net effect on groundwater quantity, but exfiltration from sewers can have a significant effect on groundwater quality (Chisala and Lerner 2008; Lerner 2002).

In Spain, the Spanish Geological Survey (IGME) carried out a study to assess the influence on groundwater quality of both the accidental spills in sewage nets and the spills coming from industries located inside the municipalities (De la Orden Gómez 2006).

Urbanization in some coastal areas with recreational facilities, along the Mediterranean coast for example, has also impacted the hydrological regime and, hence, the quality of the aquifers. Recycled waste water that is used to irrigate golf courses may increase recharge and impact water quality. On the other hand, if groundwater is used for irrigation of the golf course, the total amount of permeation could diminish instead.

17.1.4 Scope of the "Groundwater-Related Aspects (Part V)"

With regard to the *water-saturated zone*, the "Groundwater-Related Aspects (Part V)" primarily focuses on the second position of groundwater mentioned in Section 17.1.2, that is, *groundwater from the perspective of a transportation mean* for contaminants, this is, as *a pathway*, and not from the perspective of a groundwater body as a protection target. In Rolle et al. (Chapter 19 of this book), a detailed description is given of the contaminant transport processes in the water-saturated soil zone, and the relevant mathematical equations and models.

Human health risks due to the consumption of contaminated drinking water originating from groundwater very much relate to the risks associated with other human exposure pathways. Therefore, these human health risks have been incorporated in the "Human Health Aspects (Part III)" the "Human Health Protection" (see Chapter 11 by Elert et al., this book). For the same reasons of compatibility, risks for the groundwater ecosystem have been described in the "Ecological Aspects (Part IV)" the "Ecological Protection" (see Chapter 14 by Posthuma and Suter, this book) from a generic perspective; Rutgers and Jensen (Chapter 15 of this book) from a site-specific perspective).

Water that is present in the *water-unsaturated* upper soil layer, that is, the pore water, is generally not considered as a protection target. However, it plays a very important role in Human Health Ecological Risk Assessment and Food Safety, since generally the contaminants available in the pore water are those most relevant for human health, soil ecology health quality of vegetables. Therefore, the Risk Assessment aspects with regard to human health risks and ecological risks in relation to the pore water are implicitly incorporated in those chapters that focus on these protection targets, namely, McLaughlin et al. (Chapter 8 of this book) and Trapp and Legind (Chapter 9 of this book) with regard to the uptake in vegetables of metals and organic contaminants, respectively, and Hodson et al. (Chapter 16 of this book) for a description of bioavailability with regard to ecological effects.

Another important aspect with regard to the water-unsaturated upper soil layer is that the great majority of groundwater enters the aquifer as infiltrating rain water and melting snow via this layer. During passage through this water-unsaturated upper soil layer, the water quality will definitely change. In case of contaminated soils, infiltrating water will be loaded with contaminants, resulting in contaminants leaching into the groundwater. In the case of clean soils, however, contaminants may be filtered out of the infiltrating water in the unsaturated soil layer. Therefore, the pore water, analogous to the groundwater, is considered to be a *means of transport for contaminants* (*pathway*) in the "Ground water-Related Aspects (Part V)" 'Groundwater-Related Risk Assessment'. Or, in other words, the pore water is considered from the perspective of contaminant leaching to the groundwater. In Mallants et al. (Chapter 18 of this book), a detailed description is given of leaching from the unsaturated upper soil layer into the saturated soil zone, along with the relevant mathematical equations.

In addition to the important roles of pore water and groundwater as pathways, the focus will also be on the groundwater as a *protection target* in the present chapter.

Note that no attention is paid in this book to the microbial contaminants in groundwater that originate from both human and animal faeces via sewer leaks, septic tanks and manure disposal, although these are of great concern for human health (e.g., Celico et al. (2004), who found several microbial pollutants, related to pasture and/or manure spreading, in different carbonate aquifers of southern

Italy). This is because most disease-causing bacteria and viruses are short lived relative to typical groundwater travel times, except in karstic aquifers or where the source is very close to the point of water abstraction. However, the parasite *cryptosporidium* travels as an environmentally hardy oocyst (thick-walled spore phase of certain protists), and has much longer survival times, thus increasing the risks from animal husbandry on surface and groundwater catchments used for water supply.

17.2 Groundwater as Protection Target

17.2.1 Human Use

The primary use of groundwater, in addition to surface water, for human beings is as a source of drinking water. Humans drink water, partially processed in other products such as sodas, coffee or tea, on a daily basis. The two biggest consuming countries in the world use 425 litres/person/day (USA) and 326 litres/person/day (Canada) (Baynes 2004). Generally, water for consumption originates from groundwater (or surface water) that has been treated in waterworks. In some cases, especially in some rural areas in third world countries, drinking water originates directly from private wells (see Fig. 17.3 for an example of a village well in rural India). Given its role as a source of drinking water, also in developed countries,

Fig. 17.3 An example of a village well, on which villagers depend for their drinking water, in rural India (photo: University of Southern Indiana, College of Nursing and Health Professions (http://health.usi.edu); reproduced with permission)

groundwater is a very important protection target. Moreover, water is used in several other domestic applications, for example, for dish and clothes washing, showering and toilet flushing. Although the water quality requirements for these applications are much less demanding, in most countries the same water supply is used for these as is used for drinking water. Only a few countries in the world – Singapore and Hong Kong, for example – have separate water supplies for toilet flushing. Anderson (2003) illustrated that water reuse in agriculture, urban areas, industry and water resource supplementation has substantial environmental benefits for Australia, such as in reducing the impact of wastewater discharges on environmental water quality. The authors also demonstrated how water reuse has economic benefits.

Groundwater bodies represent a huge reservoir for domestic water supply. In the European Union, for example, circa 70% of the piped water supply is provided from groundwater (Foster and Chilton 2003). Worldwide, this percentage equals at least 50% (Zektser and Everett 2004). In fact, with a global withdrawal rate of 600–700 km^3/year, groundwater is the world's most extracted raw material (Zektser and Everett 2004).

Using groundwater as a resource for domestic water supply has many advantages. Firstly, in most inhabited parts of the world there is a large amount of groundwater. In spite of the huge volumes withdrawn, subtracted water volumes often are readily supplemented. The second advantage of groundwater as a source for domestic water, and most certainly for drinking water, is that the upper soil layers function as a filter, physically, chemically and biologically (for the latter, see Section 13.4.3.5 on the ecological cleaning function). This *natural cleaning function* is effective and cost-efficient. Thirdly, groundwater use often brings great economic benefits per unit volume compared with surface water because of ready local availability, high drought reliability and a generally good quality requiring only minimal treatment (Burke and Moench 2000).

Groundwater is also used for agriculture and industrial purposes, in much larger quantities, in fact, than the domestic use of groundwater. Groundwater volumes used for recreational purposes (including golf courses) is relatively limited but growing.

Worldwide, clean water resources have come under increasing pressure. Since the 1960s, the worldwide demand for water has increased significantly. Gradually, more humans are requiring more domestic water per individual. Industrial and agricultural demands have also increased during the last few decades. In contrast, the amounts of clean groundwater have slowly decreased due to desiccation, salinisation, acidification, and, last but not least, groundwater contamination (due to excess nutrients from diffusely contaminated sites, and (predominantly mobile) metals and organic contaminants from point sources). To be specific, in dry regions, over-exploration of groundwater resources has resulted in severe problems of subsidence of land, and destruction of vegetation and soil life, due to the effects from extreme drought in the soil upper layers. In many inhabited regions of the world, clean groundwater is becoming scarce due to soil contamination. An interesting indicator for the increased value that clean water has is the fact that financial institutions in many countries now offer possibilities for financial investment in clean water.

Although it is beyond the scope of this book, it should be mentioned that groundwater plays an important role in energy balances, since water has a high energy storage capacity. In many parts of the world, nowadays aquifer thermal characteristics are used, mainly in terms of thermal energy storage in the summer time for the purpose of using this energy later for heating buildings in the winter time.

17.2.2 Ecological Habitat Function

Analogous to the upper soil layer, the water-saturated deeper soil layer is the habitat for many organisms, including widely differing species. Since there is no gas-filled pore phase present in aquifers – so consequently oxygen is scarce – these organisms have a totally different origin. Actually, from an ecological perspective, it is not the groundwater but the *organisms* residing in the groundwater and the corresponding *Ecological Services* that are the protection targets.

In the European Water Framework Directive (WFD) it is stated that a good ecological status of a groundwater body must take into account the requirement that the chemical quality of groundwater must not endanger in any significant way terrestrial ecosystems that depend directly on the groundwater body (European Commission 2008).

Compared to the research efforts focused on ecological processes in the upper soil, not much research has been performed on ecological processes in the groundwater. However, since the 1990s the increased interest in *Natural Attenuation* in groundwater, that is, degradation and dilution of organic contaminants as an often cost-efficient remediation technique, has led to numerous studies on groundwater ecology. As a consequence, anaerobic degradation is much better understood today.

Considering the same arguments as advanced for the protection of organisms in the top soil layer, that is, the intrinsic significance of *Biodiversity*, along with ethical, spiritual and religious principles (see Section 13.4.2), ecological protection for groundwater is essential. Also analogous to the soil compartment, the protection of *Ecosystem Services* in the groundwater is of the utmost importance from a functional point of view. The organisms in the groundwater also play a role in carbon and nutrient cycling, although their role is not as immense as it is for the soil ecosystem in the upper soil. Probably the most important function for groundwater is the so-called ‘cleaning function’ (see Section 17.5.2). Given the enormous developments in the possibilities for practical use of Natural Attenuation in the aquifer, the role of the aquifer ecosystem is highly appreciated, today.

17.2.3 Intrinsic Value

For decades, there has been an on-going, interesting discussion concerning the *intrinsic value* of groundwater. Although most humans approach the relevance of groundwater from a functional perspective, there indeed may be less rational arguments for caring about groundwater protection. From a spiritual perspective, for example, one can claim that the groundwater is a precious part of pristine nature and is essential for peace of mind, and must be respected and protected. Even more extreme, it is known that humans have brought to bear supernatural or religious arguments in the discussion of the intrinsic value of groundwater. According to Islam all water is a gift of God. In Christianity and the western world, identical traditions evolved with regard to springs such as the one at Lourdes. The words 'water' and 'wells' are often mentioned in religious readings such as the Chronicles and the Exodus. Wells are mentioned many times in the Old and New Testaments, for example, in relation to the encounter between Jesus and the Samaritan women. Water from the Zamzam well close to the Kaaba is considered holy, as it is related to the water that saved Ismail, the ancestor of the Arabs and his mother Hagar (Ezsoftech 2009). 'Living water' often gets a double meaning: as real water from a well, it is needed for any living creature to survive, and likewise the Holy Books refer to believing as the drinking of the 'living water' (E. Smidt 2009 'personal communication').

Groundwater, being less visible than surface waters, is more easily associated with holiness. It cannot be excluded that water was considered an important resource not only for religious reasons, but also because it was an extremely scarce product due to its absolute importance for human life and its scarcity in arid areas such as Palestine. A similar hypothesis could contribute to the explanation of the worshipping of wells in several cultures, generally in regions where water is a scarce resource.

Taking into account religious and spiritual wisdom in water management can play a positive role in promoting wise water management by linking values and actions that promote sustainable water management (Schelwald-Van der Kley and Reijerker 2009).

17.2.4 Sustainability

One of the biggest challenges with regard to a sustainable environmental policy is found in groundwater protection. In general, sustainability relates to a specific state of a process, or state, that is able to continue or to be maintained indefinitely. In this book a wider definition such as the Brundtland Commission definition of the General Assembly of the United Nations (United Nations 1987) is used (see Section 1.9.6): a development that meets the needs of the present without compromising the ability of future generations to meet their own needs *or negatively affecting the wider surroundings*. This expansion of the common sustainability definition is especially relevant with regard to groundwater as a pathway for contaminant transport, since contaminant transport may lead to risks at places far from the source.

Several groundwater reservoirs have been used as sustainable drinking water resources for many decades, or even centuries. As a consequence, the general belief often was that at least the next few generations might be able to benefit from the same groundwater reservoirs. In the past, a sustainable water supply was typically considered to be a matter of water quantity, in other words, that it was related to groundwater extraction and replenishment, and was believed to be guaranteed through supplementation by infiltrating rain water to counterbalance groundwater extraction. During the last few decades, however, many of these groundwater reservoirs have been threatened by groundwater contamination, which has provided a new dimension to sustainable water supply. To ensure the drinking water needs for future generations, it is necessary to prevent contaminants from the upper soil layers from leaching into the groundwater and, hence, transporting vertically into aquifers and, indirectly, into surface water reservoirs. These processes might take decades, and sometimes centuries. To ensure drinking water supply at locations other than at contaminated sites, lateral transport of contaminants with and within the groundwater must be controlled.

Vrba and Lipponen (2007) derive ten groundwater indicators for assessing the sustainability of specific groundwater bodies, including an indicator for soil quality, based on the area percentage of an aquifer in which the concentration of the indicator parameter exceeds the maximum level specified in the WHO drinking water guidelines (or equivalent).

17.2.5 Appreciation

17.2.5.1 General Public

Water is generally recognized by the general public as a precious material, although compared to other precious materials (e.g., gold, diamonds, or oil) water is ubiquitous and available in large amounts. Some extreme statements about the meaning of water have been heard such as ‘water promises to be to the 21st century what oil was to the 20th: the precious commodity that determines the wealth of nations’ or ‘water scarcity is now the single biggest threat to global food security’ (Optimum Population Trust 2009).

Since early mankind, an essential criterion for pioneers to create settlements was the availability of reliable water resources. Water, in its meaning as the liquid product that we drink, generates an extremely positive sentiment in humans. Many pictures in magazines and on websites, for example in advertisement of mineral water, or with regard to sports or health programs, show young healthy looking people in association with drinking water (see Fig. 17.4 for an example). It is commonly associated with life, life in the widest sense of the word: the life of plants and organisms and, last but not least, human life. More than half the human body is made up of water, while blood contains as much as 95% water. Quite simply, almost every living organism needs water to survive. Humans do not have to be experts, that is, (geo)hydrologists, to understand the absolute importance of water.

From the perspective of the wider meaning of water, there is also another aspect to the appreciation of water. This type of appreciation strongly depends on the climate. In countries with a wet climate people may complain about the many rainy days, while in countries with a dry climate rainy days may be celebrated. Water, as in huge volumes like lakes, rivers and oceans, is also feared for its destructive

Fig. 17.4 An example of a picture showing young healthy looking people in association with drinking water, illustrating the extremely positive sentiment in humans that water, in its meaning as the liquid product that we drink, generates (photo RIVM; reproduced with permission)

potential due to flooding. And in many parts of the world where water has traditionally been available in seemingly endless amounts, this ubiquitousness has led, and leads today, to large-scale spilling.

However, in reference to water, few people think about *ground*water. Generally speaking, if groundwater is ever appreciated by a large part of the general public, it is for one single reason: it gives us drinking water. And, one thing is certain: those individuals who are dependent on private water wells are very much aware of the valuable liquid that resides in the subsoil. One interesting aspect of all this is the widespread appreciation for mineral waters, that is, waters originating from entities of the earth's crust that have not been influenced by human activities and generally are only influenced by soil and rocky materials from which the water takes its minerals. These waters usually have had a relatively long residence time, and, hence, have had more time for minerals to dissolve in them. These mineral waters are generally associated with a healthy lifestyle and sometimes relate to a 'back to nature' kind of attitude. In fact, mineral water comes from the same source as drinking water from the tap, namely groundwater, but from locations deeper or further away from human influences, so that purification measures generally are not necessary. It can happen, however, that drinking water limits are exceeded in mineral water, for example, with regard to sulphate concentration.

Typically, groundwater is one of the few natural resources that are renewable within a short time frame, that is, in relation to the abstraction rates by humans. However, since contaminants can migrate easily from the surface down to the groundwater and are then contained within the pore spaces of the underground matrix, they are not always easy to eliminate. Groundwater is therefore very vulnerable to contamination.

The general public is only barely aware of the fact that contamination of water resources occurs; some more interested individuals may have read about this in the newspapers or learned about it via television. And some people then use this information to criticise the pressure on the environment in general (‘one cannot even trust the groundwater that comes out of the tap these days’).

17.2.5.2 The Groundwater Ecosystem

The ecosystem in groundwater is even more underestimated than the ecosystem in the upper soil layer (see Section 13.1.3). Many humans will be surprised when they learn that such an ecosystem in groundwater even exists. Some non-experts might have read articles in the newspaper about the surprising abilities of organisms in the groundwater to remove (in reality, break down) harmful organic contaminants. This group of people generally is interested in, if not overwhelmed by, the phenomenon of natural contaminant removal. However, they usually do not have any idea about the magnitude of Natural Attenuation processes and the costs that are involved when human beings have to perform these same tasks using extraction and purification technologies.

17.2.5.3 Political

Politically, groundwater was considered to be one of the precious resources long before contaminated sites came onto the political agenda. All developed countries have had legislation for many decades on groundwater protection. However, originally this legislation was primarily initiated from the perspective of the protection of drinking water resources from a quantitative point of view. During the last few decades, water quality issues have gained in importance.

Groundwater is a means of transport (carrier) for contaminants, that is, it is a pathway that will lead to effects at other locations. The most important aspects of the groundwater pathway are the leaching of contaminants from the upper soil to the deeper soil layers and subsequent groundwater flow in the water-saturated soil layer. Both these aspects play a major role in Risk Assessment. From this perspective, in addition to groundwater as a protection target, groundwater as a contaminant carrier is an implicit part of most policies for contaminated sites management.

For decision-makers and regulators, the protection of the groundwater ecosystem has long been a difficult issue. Both in terms of Biodiversity as well as in terms of Ecosystem Services, the groundwater ecosystem really is not well understood. The significance of Biodiversity in the groundwater is still a matter of debate, since the species diversity, although huge, is not as high as in the upper soil. Moreover, these organisms are even more hidden than their terrestrial counterparts are. This is different vis-à-vis the significance of the Ecosystem Services of the groundwater, mainly in terms of its ‘cleaning function’. Since in situ remediation technologies offer extremely important effective solutions, Natural Attenuation is a key process in groundwater Risk Management. For this reason, no decision-maker or stakeholder doubts the significance of the abilities of the groundwater ecosystem today.

From this perspective it is interesting to note that appreciation for the soil and groundwater ecosystems has undergone a dramatic evolution among ‘problem owners’. A few decades ago these ‘problem owners’ were confronted with the need to protect the soil and groundwater ecosystem whose intrinsic value from their point of view hardly seemed to counterbalance the high remediation costs. But today these stakeholders make excellent use of these ecosystems by using them to degrade the contaminants that they must be protected against.

17.3 Groundwater as Contaminant Pathway

17.3.1 Source-Pathway-Receptor Approach

With regard to the *source-pathway-receptor approach*, groundwater provides an important pathway for contaminants, connecting a source to one or more receptors. The transport velocity of water and contaminants is strongly dependent on soil type, varying from high speeds in gravel and coarse sand to much slower in fine sand and clay soils. Moreover, since contaminants undergo an intensive exchange with soil materials (adsorption), the transport velocity of contaminants strongly depends on the contaminant characteristics and adsorption properties of the soil and aquifer material. Generally, contaminants show a high adsorption affinity towards humic acids, organic clays and oxides covered with organic coatings (McCullough et al. 1999). The principles of transport significantly differ for the water-unsaturated upper soil versus the water-saturated groundwater zone. Generally speaking, the transport of water and contaminants is much faster in the groundwater zone than in the water-unsaturated upper soil layer. It should be noted, however, that surface water in rivers generally moves much faster than the groundwater.

Obviously, the type of receptors (protection targets) that are threatened depends on the direction of groundwater flow. Human health may be affected when contaminants migrate downwards and reach drinking water extraction points. Less frequently, human health may be threatened when humans come in contact with contaminants after an upwards flow of groundwater. A major pathway for human exposure is through inhalation of contaminated indoor air, and to a lesser extent outdoor air, after vapour intrusion of contaminants from groundwater into buildings in residential, industrial, or recreational areas. This pathway is intensively described in McAlary et al. (Chapter 10 of this book).

Ecological protection targets are reached after the leaching of pore water into the groundwater, since the aquifer ecosystem will then experience a subsequent exposure to contaminants. Indirectly, contaminated groundwater may impact terrestrial ecosystems in nature reserves and agricultural products after an upwards flow. Eventually, contaminated groundwater will threaten the surface water ecosystem, and eventually in some cases the oceans, after drainage into drain tubes, ditches, canals and rivers.

When an intrinsic value is attributed to groundwater (see Section 17.2.3), ‘clean groundwater adjoining a contaminated groundwater body’ also should be considered as a protection target. In that case, the two different meanings of groundwater in Risk Assessment, that is, as a protection target and as a means of transport (carrier) for contaminants, come together: migrating contaminated groundwater (the pathway) contaminates clean groundwater (the protection target).

Several types of measurement for risks from groundwater transport exist. With regard to a threat to identified protection targets, often some specific measurements are used such as:

- the estimated concentration in the layers where groundwater is extracted for drinking water consumption (protection target: human health);
- the estimated concentration in surface water (protection target: aquatic ecological effects);
- the estimated concentration in the root zone after an upwards flow of groundwater (protection target: Food Safety).

With regard to the intrinsic value of groundwater, the volume of clean groundwater in which a specific risk level threatens to be exceeded within a specified timeframe is often used as measure for risks. The more general measurements used are contaminant profiles with depth at several specified moments or contaminant concentration profiles over time at a specific depth (*breakthrough curves*).

17.3.2 Transport Characteristics

17.3.2.1 General Transport Pattern

The transport in the water-unsaturated upper soil is predominantly vertical. The velocity of water flow and, hence, contaminant leaching into the groundwater is strongly dependent on the water content of the upper soil. This velocity is minimal when soil is at its driest level, gradually increasing as soil becomes wetter. The maximum water flow velocity in the water-unsaturated upper soil, under (temporary) water-saturated conditions, equals the velocity of groundwater flow in the saturated zone.

Groundwater flow in the water-saturated soil layer is the major contaminant transport mechanism in soils, bringing contaminants from as far away as many kilometres from the source. The transport direction in the water-saturated soil layer is predominantly horizontal. As a consequence of this contaminant transport, contaminants may threaten just about any protection target.

The sources of and pathways for groundwater recharge in urban areas are more numerous and complex than in rural environments (Lerner 2002). Buildings, roads, and other surface infrastructures, combined with man-made drainage networks, change the pathways for precipitation. Some direct recharge is lost, but additional recharge can occur from storm drainage systems. Large amounts of water are imported into most cities for their water supply, distributed through underground pipes, and then collected again in sewers or septic tanks.

17.3.2.2 Impact of Heterogeneous Soils or Aquifer

The subsurface of the soil is generally structured in layers. However, these layers do not follow a complete geometrically symmetrical pattern, but show significant irregularities. Moreover, in urban areas the soil layers are pierced with many civil constructions. And some soils show a fissured structure. Consequently, groundwater flow and contaminant transport show an irregular pattern.

17.3.2.3 Impact of Surface Water Bodies and Anthropogenic Subsurface Processes and Structures

The presence of *surface water bodies* such as canals, rivers, lakes, ditches, etc., does impact the groundwater flow and contaminant transport, since the flow lines tend to move towards these groundwater bodies. Generally speaking, the bigger the groundwater body, the higher the impact. Additionally, many *anthropogenic subsurface processes and structures* influence groundwater flow and contaminant transport. A typical example of an anthropogenic impact on the groundwater flow is *the presence of public groundwater extraction wells*, often at depths of several hundred meters, or of *private wells*. Obviously, these wells impact the groundwater flow and contaminant transport pattern. Many *civil structures* intrude into the subsurface, mainly in urban areas, and so impact the groundwater flow and contaminant transport. Examples of such structures are subsurface car parks, tunnels, groundwater wells, water supply nets, and sewage pipelines.

Other examples of anthropogenic impact are *hydro modification projects*, which have a substantial effect on the natural regime of surface water and, indirectly, groundwater. The construction and operation of dams, for example, may magnify the transport of contaminants into the groundwater body due to the higher hydraulic pressure differences. The construction of a dam or a reservoir can modify the relationship between surface water and groundwater, and therefore may result in increased recharge and discharge. River regulation using reservoirs, channel impermeabilisation of natural streams, canal and irrigation channels, etc., are some other factors that impact water flow and contaminant transport.

Underground or open cast mines, quarries and mineral extraction activities also strongly impact the flow pattern, since large withdrawals of groundwater take place when dewatering a mine in preparation for mining activities. Moreover, if a deep mine intersects an aquifer, variations in the natural flow regime and changes in recharge or discharge may result.

A specific case of anthropogenic impact on the groundwater flow, and hence of contaminant transport, is found in *polder areas*, that is, isolated areas with an artificial groundwater table deeper then the groundwater table in the surroundings, and often below sea level. Characteristically, the groundwater level in these polders is manipulated by pumping water out of the polder area into surrounding canals. These polders are generally found in coastal areas and are most common in the Netherlands.

Irrigation may also increase water recharge significantly. In the past few years, new irrigation areas have been created in many parts of the world. They have contributed to an increment of groundwater recharge. Clearing of native vegetation in recharge areas, as in upper parts of a catchment (which usually correspond to the lowest agricultural productivity areas), can increase the amount of recharge as well. One side effect of such an increase in recharge can be that salt and contaminants in the upper soil layers are leached into underlying aquifers, causing land salinity and groundwater contamination.

Construction of roads can affect both recharge and discharge. If the road is in a recharge area, contaminants may be released accidentally during the construction or operation phases. In regions with colder climates, salt added to melt snow in winter can leach into the groundwater. When the road is in a discharge area, contaminants are washed into the surface water. Movement of equipment and supplies can have an impact as well.

17.4 Calculating Contaminant Transport

17.4.1 Conceptual Models

Every appropriate model calculation must be preceded by the development of a Conceptual Model. In Section 1.8.7.2 a Conceptual Model was defined as a (usually visual) presentation of the relationship between the source, all the pathways involved and the receptor. A cross-section of the contaminated site is the most common format of a Conceptual Model. An example of a Conceptual Model, as a cross-section of a groundwater body, is given in Fig. 17.5 (European Commission 2008). In this graph, sources, pathways and receptors are given. Moreover, several *Points Of Compliance* (*POC*) are identified in the graph in the framework of the Water Framework Directive, which help to determine whether a contaminated discharge is acceptable or to decide upon required remediation levels (European Commission 2008). POC 0 relate to the concentrations directly near the source in the unsaturated zone. The groundwater at POC 1 to POC3 at different locations in the saturated zone could be considered as receptor. There are two types of compliance point, these are, a theoretical point within a model for calculating an acceptable discharge concentration or required remediation level, and a physical monitoring point (e.g., an observation borehole). A compliance point may either be at the receptor itself or, for practical reasons, at a point between the source and the receptor.

The development of a Conceptual Model is essential for modelling contaminant transport, as groundwater bodies are complex systems, and the Conceptual Model helps in understanding and identifying the paths, receptors and points (or planes) of compliance. Moreover, it provides a schematisation of the geometry of an aquifer. In addition, a Conceptual Model provides support for explaining Groundwater-related Risk Assessment to the stakeholders involved.

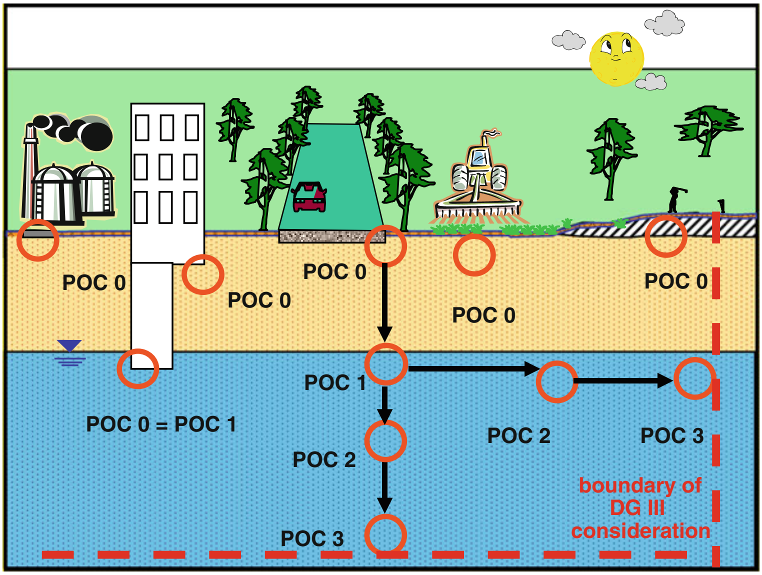


Fig. 17.5 A cross-section of a groundwater body, as a Conceptual Model with regard to contaminant transport (European Commission 2008, reproduced with permission)

After the first draft of the Conceptual Model has been constructed, the first aspect to be considered is the geometry of the geologic formations (Grima et al. 2009). In the real world, groundwater bodies can range from simple aquifer formations to very complex multilayer aquifer-aquitard systems. There is a wide variety in the range of geometries, from a uniform morphology with limited spatial variation (such as tabular formations) to complex formations where processes of deformation and fracturing of previous structures are dominant. Hydrodynamic parameters, such as hydraulic conductivity and storage, will affect the flow pattern and the hydraulic gradient and, consequently, the potentiometric surface and its evolution. The flow regime is laminar in detritic aquifers, while in fractured rock aquifers the flow is faster through preferential channels, and turbulent flow can take place when channels have a given proportion. The hydrostatic pressure of groundwater within an aquifer is another important element that needs to be accounted for. In unconfined aquifers, hydrostatic pressure is represented by the groundwater table, whereas in confined aquifers the water level will depend upon the pressure head. Subsequently, three basic aspects of aquifers must be dealt with (Grima et al. 2009): aquifer parameters, pressure head regime and the hydrogeochemical features of groundwater.

17.4.2 Mathematical Models

17.4.2.1 Role and Principles

Mathematical models have at least two separate roles. Firstly, they support the understanding of the groundwater flow and contaminant transport system. By testing hypotheses and using an iterative approach to modify the model in the light of new data or improved understanding, the practitioner reaches the best representation of the system being studied. Secondly, models support decision-making. It is generally recognized, however, that models are not perfect nor ever will be, in representing complex systems such as the subsurface environment.

Basically, contaminant transport includes four types of processes; these are water flow, migration of contaminants within the water phase, physico-chemical interaction of contaminants in the pore water or groundwater with the soil material, and degradation. For many decades, at least since the famous French scientist Henry Darcy formulated his simple groundwater flow equation in the mid of the 19th century (Darcy 1856), engineers have been able to estimate the water flow velocity. Today, many equations and a wide variety of models exist to calculate water flow and contaminant transport.

Generally, the calculation of saturated groundwater flow is equivalent to calculating electrical current: the flow is proportional to the hydraulic gradient (as compared with the gradient in voltage) and inversely proportional to the hydraulic resistance (as compared with electrical resistance). Groundwater flow in the unsaturated zone (pore water), however, is far from linear. This makes the calculation of contaminant transport in the water-unsaturated upper soil layer much more complex than in the saturated groundwater zone. However, analytical solutions for contaminant transport in the water-unsaturated upper soil layer, like those for the saturated zone, exist for a whole series of initial and boundary conditions.

17.4.2.2 Numerical Models

Since the early 1970s, groundwater engineers have used computers to overcome the difficult problem of non-linearity. Today, numerous *numerical models* for estimating the groundwater flow and contaminant transport in both the saturated and the unsaturated upper soil layer are available for every type of contaminated groundwater problem. Such a numerical model calculates the water flow and contaminant transport between concrete soil layers in a series of concrete time steps, while the calculations in each time step build on the results from the former time step. To this end, for one-dimensional numerical models, for example, a contaminated site soil profile must be subdivided into hypothetical concrete layers with fixed thickness (often a few centimetres), or, in case of more complex models into concrete layers with different thicknesses. Time steps can vary from seconds to days, depending on the complexity of the transport problem. An example of the use of numerical models is given in Chang et al. (2001), who used a numerical model to calculate copper and cadmium transport in a lateritic silty-clay soil column.

Some numerical models include the possibility of assessing the transport of multiple contaminants. An example of such a model is given in Mayer et al. (2002), who calculated the transport of organic contaminants in an unconfined aquifer overlaid by unsaturated sediments and of acid mine drainage in the unsaturated zone of a tailings impoundment at the Nickel Rim Mine Site near Sudbury, Ontario, Canada, and the subsequent reactive transport in the saturated portion of the tailings.

Even more complex than the presence of multiple contaminants is the presence of liquids with a different density in the groundwater. Light Non-aqueous Phase Liquids (*LNAPLs*), for example, are lighter than water and may often float on the water table. Dense Non-aqueous Phase Liquids (*DNAPLs*), on the contrary, have a higher weight than water. Density differences may lead to so-called stable or unstable displacement. The instability of the flow pattern will increase, and show a more chaotic behaviour, when the density differences are bigger. Also fluid density variations caused by changes in the contaminant or colloidal concentration, temperature, or pressure of the groundwater, may drive contaminant transport (e.g., Simmons et al. 2001). Except for transport of pure liquids, the LNAPLs and DNAPLs are a source for contaminants which may migrate from the pure liquid into the adjoining (clean) water bodies. An example of a transport model that calculates the transport of contaminants originating from the dissolution of DNAPL pools in aquifers is given in Tatalovich et al. (2000).

The following processes control contaminant transport in a more or less homogeneous soil:

- convection (advection): migration of a mass of groundwater (and soil gas in a water-unsaturated upper soil layers), driven by pressure differences;
- (molecular) diffusion: migration of contaminants within the groundwater (and, when present, soil gas), driven by differences in contaminant concentrations;
- (hydrodynamic) dispersion: migration of contaminants within the groundwater, driven by irregularities of the soil pores and of the flow velocity pattern within a pore;
- adsorption and desorption: exchange of contaminants between the solid phase of the soil and the pore water or groundwater;
- degradation: the biological or chemical reduction of a contaminant to one or more less complex compounds;
- water uptake by roots (upper unsaturated layer, only).

As an alternative for adsorption and desorption processes, the chemical part of the models may be described on the basis of thermodynamics as, for example, shown in Meinrath and May (2002) with regard to contaminant leaching from mine waste.

17.4.3 Reliability of Model Calculations

17.4.3.1 Uncertainties

For many decades, there has been a debate on the reliability of models that predict contaminant transport on the basis of the processes mentioned in Section 17.4.2.2. In fact, this debate is still ongoing. The reliability of groundwater flow and contaminant transport models is a complex issue, even in homogeneous porous media. An abstraction of the reality must be done to be able to use models.

In spite of the sophisticated mathematical background, a numerical model based on a regular homogeneous porous medium is often insufficient. Examples of such cases are heterogeneous soils or aquifers (Section 17.3.2.2) and the presence of surface water bodies and of anthropogenic subsurface processes and structures in the underground (Section 17.3.2.3). It is a challenge, but very important, to include these factors in numerical models.

Homogeneity and heterogeneity of a system are concepts closely related to the scale of the investigation. The same groundwater body can be considered as homogeneous on a large scale and heterogeneous on a small scale. Therefore, the scale of investigation must be considered at the very beginning of a project. This scale strongly depends on the relevant receptor, and is different, for example, for adjacent clean groundwater and public water supply groundwater bodies a few kilometres away. In case of the development of a Risk Management solution, the relevant scale depends on the type of solution, and differs, for example, for source removal and pathway interception.

In Section 17.4.1, the importance of a Conceptual Model with regard to appropriate modelling of contaminant transport has been outlined. A correct Conceptual Model could significantly increase the reliability of contaminant transport modelling. Findings based on a wrong Conceptual Model, or on no Conceptual Model at all, can produce results in a direction contrary to what would normally be expected.

It is generally recognized, however, that relatively large uncertainties are involved when using models to assess risks from contaminant transport. The reason for this is that usually drastic simplifications have to be adapted. These simplifications relate to the subsurface geometry, hydraulic and geochemical input parameters, and the magnitude and type of the source.

Subsurface Geometry

The subsurface geometry involves huge volumes of subsoil: plumes generally move through volumes of soil of several hundreds to thousands of cubic meters within the time span of one year. Although these large volumes often are made up of several irregular layers from different materials, and often include entities of another different subsoil material, the use of contaminant transport models requires the design of a more or less regular profile. An important question in the debate on reliability of contaminant transport calculations is: can a layered soil at the relevant scale, with some cracks and holes be considered as a ‘regular’ (homogeneous) porous medium? Besides this, it is hard to identify locations in the subsoil that allow for very fast, preferential water flow and contaminant transport, for example, due to cracks or soil materials with high hydraulic permeability. Fissured soils are difficult for accurate contaminant transport calculations. In these soils, water and contaminants follow preferential flow paths. Bruell and Inyang (2000) described the difficulties when modelling contaminant transport with regard to remediation of groundwater in contaminated rock masses. Failure to account for preferential flow paths in numerical simulation can lead to over-estimation of the effectiveness of the remedial measure under consideration (Zheng and Gorelick 2005).

The success of groundwater transport estimates depends on the schematization of irregular features in numerical models. MacQuarrie and Mayer (2005), for example, modelled contaminant transport in fractured crystalline rock environments as discrete fracture or dual continuum media. Several other models include preferential flow due to cracks or soil materials with a high hydraulic permeability. Zheng and Gorelick (2005), for example, included this phenomenon in their model by decimetre-scale preferential flowpaths. They used a hypothetical but geologically plausible network of 10 cm wide channels of high hydraulic conductivity, as a representation of the relative preferential flow paths. Field and Pinsky (2000) used a two-region non-equilibrium model (including partitioning of solute into mobile- and immobile-fluid regions) for the calculation of contaminant transport in karstic aquifers.

Moreover, the impact of the factors described in Sections 17.3.2.2 and 17.3.2.3 on contaminant transport, which are the presence of heterogeneous aquifers, surface water bodies and anthropogenic subsurface processes and structures, is difficult but essential to quantify. This phenomenon results in an uncertainty that is difficult to assess, and which implies a challenge for practitioners.

Hydraulic and Geochemical Input Parameters

With regard to hydraulic conditions, a specific hydraulic gradient has to be assumed over relatively large areas. Geochemical input parameters have to be determined for large volumes of aquifer material, often of an irregular nature. As a consequence, the description of both hydraulic and geochemical input parameters contributes to uncertainties in the predicted groundwater quality. It is a well-known fact that, as for any other model, the availability of adequate input parameters is crucial for the quality of the model calculations. Several guidelines exist for contaminant transport model parameter identification (e.g., Lovanh et al. 2000). With regard to the parameter identification, it is a huge advantage when the contaminant transport modeller is familiar with the geometry of the soil and aquifer.

Interaction processes between water-bearing soils and rocks and the water itself generate different water quality sectors, and this fact will determine the natural evolution of every particular system. The effect of interactions with hydraulically connected aquifers as well as surface water needs to be considered when looking for a detailed description of heterogeneities. This factor is particularly important when dealing with coastal aquifers, where a freshwater and saltwater interface exists. Interactions between both types of waters and geological formations must be accounted for, since seawater intrusion is a dynamic and three-dimensional process that causes water quality variations at horizontal and vertical scales.

The processes to which calculated contaminant profiles with depth or contaminant breakthrough curves are most sensitive depends on the contaminant characteristics. Mobile contaminants, for example, are generally most sensitive to water flow-related processes and input parameters, while immobile contaminants are more sensitive to adsorption and adsorption-related input parameters (Swartjes et al. 1993).

Source

One of the most uncertain factors in contaminant transport is the quantification of the source, that is, the emission to soil and groundwater. Examples of uncertainties with regard to the source relate to the necessary simplifications of the irregular contaminant inputs. In models, these inputs must be translated into a homogenous contaminant load in a specified layer of the (upper) soil or into a steady-state contaminant input at the surface of the soil or into the water-saturated soil compartment.

17.4.3.2 Dealing with Uncertainties

Most of the issues described in Section 1.5.4.2 on dealing with uncertainties from a general Risk Assessment perspective, also apply to Groundwater-related Risk Assessment. An important conclusion is that outputs from contaminant transport models must be regarded as indications. Nevertheless, these models are extremely useful tools, as long as they are intelligently used. A few recommendations for the use of contaminant transport models are:

- relatively simple worst-case calculations can be used in a first tier Risk Assessment;
- generally, trends can be assessed in terms of a protection target that might be threatened by contaminant transport within a specific time range (e.g., ‘within a few years’);
- identification of the most sensitive processes and input parameters supports good modelling practice;
- in specific cases the modeller could consider performing a probabilistic model calculation;
- in sensitive cases peer review or second opinions may be established;
- the uncertainties must be made transparent and must be communicated to all stakeholders;
- methods can be indicated to reduce the uncertainties during subsequent project phases.

17.4.4 Good Modelling Practice

As with any model, for the reasons previously mentioned, no existing contaminant transport model can ensure the correct prediction of the development of transport modelling, that is, the evolution of groundwater plumes or the breakthrough curves of a contaminant. Any practitioner should realise this when applying models for the secondary purpose of contaminant transport modelling as described in Section 17.4.2.1, that is, understanding contaminant transport systems and supporting decision making on the basis of calculated concentrations.

Modelling of contaminant transport is especially complex when several groundwater plumes interact, a situation that is often found in intensively urbanised areas.

For a successful calculation of contaminant transport, the following is needed:

- a detailed understanding of the geometry of the subsoil;
- an understanding of the source;
- a thorough understanding of the processes included in the contaminant transport model along with the algorithms to describe these processes in terms of hydraulic and geochemical processes;
- an understanding of the range of applicability of the contaminant transport model;
- an understanding of the sensitivity of the input parameters;
- a proper assessment of (the most relevant) hydraulic and geochemical input parameters;
- an awareness of the (lack of) reliability of the model calculations.

In other words, a contaminant transport modeller must not only understand the software package, but must have a thorough understanding of the context of the model application (the Conceptual Model) and the scientific background of the model and the processes involved. This requirement is challenged at a high level these days, since many user-friendly numerical contaminant transport models are, often freely, available on the internet.

Obviously, experience in contaminant transport modelling and the intuitive skills of contaminant hydrogeologists support good modelling practice. Since a decent model does not a priori give good model results, and in order to offer the possibility of verification, the model user is obliged to justify the modelling process.

A serious mistake that must be avoided is distinguishing between validated and non-validated models. The phrase ‘the myth of the validated model’ could be the title of a suspense novel but is, in the context of this chapter, a serious warning for model misuse, since, in fact, not models but only *model applications* can be validated. Validated models hardly (if ever) exist (Leijnse and Hassanizadeh 1994). Therefore, the term ‘validated model’ must not be used as a quality mark to convince non-experts of the performance of complex models. Although a series of validated model applications might increase the reliability of the model, there are so many conditions that could differ for other model applications (e.g., aquifer type, degree of heterogeneity, concentration range and presence of ‘disturbing factors’ such as groundwater extractions or surface water bodies) that the unconditional reliability of the transport calculations with the same model is nothing but an illusion.

The choice of model very much depends on the purpose of the model activity, for example, the necessity of a rough estimate of contaminant transport velocity or the need for a more or less precise contaminant depth profile, a local or diffuse contaminant pattern, etc. The choice of model also depends on the type of Risk Assessment. In a first-tier Risk Assessment, for example, a simple conservative model calculation could be sufficient. When this exercise results in the exclusion of unacceptable risks, the Risk Assessment is concluded and usually no further action has to be taken. When unacceptable risks cannot be excluded, the use of a more complex model may be appropriate in a higher tier. These models, in higher-tier Risk Assessments, often are numerical models.

17.5 Risk Management

17.5.1 Scope

In the Introduction chapter (Section 1.6.1) Risk Management was defined as the design and application of strategies for dealing with risks. Risk Management is appropriate when the conclusion from a Risk Assessment is that a particular risk is unacceptable. It includes avoiding the risks, mitigating or removing risks and, last but not least, communication about the risks with the parties involved. It was concluded in this section that the keyword in Risk Management is *risk reduction*. Other conclusions from this section, which also apply to contaminated groundwater, are:

- ways to achieve risk reduction are source control treatment, or blocking the pathway from source to receptor;
- the challenge is to find the optimum balance between the most effective and most cost-efficient way of doing this by weighing the short-term advantages and the costs of aftercare.

It also was stated in the Introduction chapter that remediation, that is, in its most strict definition elimination of the source and the resultant groundwater contamination, is the most direct way of risk reduction. However, remediation often is too drastic an activity, whose results are not in alliance with the social and technical impact at the site and the costs.

Risk Management approaches for contaminated groundwater are still evolving. In the late 1970s, Risk Management was often the same thing as complete removal of the contaminants and, hence, of the risks involved. *Pump-and-Treat* was the most popular mechanism to achieve this goal for groundwater contamination. Since groundwater is difficult to assess, however, groundwater remediation is relatively expensive. Moreover, given the complex exchange of contaminants in the groundwater and the solid phase of the aquifer, the course and time lapse of Pump-and-Treat remediations are difficult to predict.

Today, many creative solutions for the Risk Management of groundwater are applied. Nevertheless, the prevention of groundwater contamination (aka: *prevent and limit approach*) is by far the most cost-efficient approach with regard to the maintenance of a sustainable groundwater quality.

17.5.2 Natural Attenuation

Modern groundwater Risk Management is nearly inconceivable without the cleaning qualities of the organisms. Since the mid 1990s *Natural Attenuation*, that is, degradation and dilution of contaminants in the groundwater, evolved into an extremely important instrument in groundwater Risk Management. Natural Attenuation procedures are often combined with Pump-and-Treat technologies, nowadays for a detailed exposition on Natural Attenuation, see Chapter 22 by Peter et al., this book.

17.5.3 Regional Approaches

It is recognised that it is much more cost-efficient to simultaneously remediate a series of groundwater plumes in a region than to do this site by site, certainly in the case of intermingling groundwater plumes (see Section 17.7.2 on mingling groundwater plumes). Moreover, it is advisable to manage groundwater and surface water as an integrated whole (Environment Agency 2009).

To be able to bring this *regionally based groundwater Risk Management* into practice some tough judicial problems have to be overcome. Firstly, the responsible parties who caused the different plumes need to be identified. This is sometimes difficult, especially in the case of intermingling groundwater plumes that were caused a long time ago. Secondly, depending on the responsible parties identified, the contribution to the total costs of the different stakeholders has to be negotiated. Given the complexity of the legal and technical situation and the costs involved with regionally-based groundwater remediation, hydrogeologically trained mediators can play an important role in this process.

17.6 Sampling and Monitoring

17.6.1 Purpose

The most common purpose of groundwater *sampling* is to gain insight into the actual contamination pattern in the groundwater. Besides this, time series of concentrations of contaminants in groundwater, assessed through *monitoring*, are important in evaluating the developments in groundwater quality. Monitoring networks are also constructed for feeding data to models that allow simulation of groundwater flow and contaminant transport. In Fig. 17.6 an example of a groundwater sampling activity at a site in Ft. McMurray, Canada, is given.

Sampling and modelling should be intimately linked in an iterative way. Data feed models which predict outcomes. If the outcomes are not validated, then often the data are deficient, and so the data input needs to be modified.

Sampling or monitoring objectives must be defined at the early stages of a contaminated site investigation.

17.6.2 Groundwater Concentration Pattern

Generally, the spatial variation in groundwater concentrations is less than the spatial variation in soil concentrations due to the mixing of groundwater and contaminants in groundwater. Nevertheless, accurate measurements and monitoring are burdened with several difficulties. For bigger sites, for example, the spatial representation of samples is often limited, since groundwater sampling is relatively expensive. Even when a high density of sampling wells is applied, the total surface of the wells generally is only a small part of the area under investigation. Besides this, problems

Fig. 17.6 An example of a groundwater sampling activity at an industrial site in Alberta, Canada (photo: courtesy of Alberta Innovates – Technology Futures; reproduced with permission)

of poor construction and maintenance of wells can cause local contamination at the well, which is not necessarily related to the contamination of the aquifer (Vrba and Lipponen 2007). From these restrictions it can be concluded that it is difficult to get a detailed picture of the contaminant pattern in groundwater.

At most contaminated sites, only the upper groundwater is sampled. This is simpler and cheaper, and for most Risk Assessments the quality of the upper groundwater is the most relevant. For deeper groundwater quality, the risk assessor is often dependent on data from public-production wells, although data on raw water quality may be difficult to obtain from the waterworks. Moreover, in deeper wells water from different origins may have been mixed.

In Lamé (Chapter 3 of this book) the procedure for sampling has been described primarily from a practical perspective. In Brus (Chapter 4 of this book) this procedure is approached from a statistical perspective.

17.6.3 Lysimeters and Column Experiments

Contaminant transport cannot be measured directly in the field. It is possible, however, to monitor groundwater concentrations which indirectly relate to contaminant transport. There are several other experimental procedures that are used to assess the groundwater flow and contaminant transport, the most common being lysimeter and column experiments. Lysimeters are reservoirs which were originally used for the assessment of water fluxes through soil, for example, by measuring evapotranspiraton rates. In a similar way contaminant fluxes can be measured within lysimeters, either in the field or in the laboratory. As a measurement for contaminant transport, it is possible to monitor the leachate that drains from the lysimeter either as concentration over time or as the accumulated amount of leached liquid and contaminants. An example of a lysimeter study is found in Kale et al. (2001), who studied the distribution of PCBs (polychlorobiphenyls) (congeners 28 and 52) and PAHs (fluoranthene and benzo[*a*]pyrene) in German soils.

Column experiments are performed in the laboratory. For this purpose, a metal or plastic column is filled with soil material. The soil material then preferably mimics the natural setting of the soil as much as possible. That means that for the most optimal column experiments a column of soil is cut out from the soil. An example of a column experiment is given in Mayes et al. (2000), who investigated the hydrological and geochemical processes that govern the fate and transport of, among others, metals in undisturbed saturated columns of weathered, fractured shale saprolite.

Column experiments have the major advantage of enabling full control of the transport conditions such as the water and contaminant input on the surface of the soil material in the column. The major disadvantage is that it is often hard to maintain the layered soil structure of natural soils in a column. Moreover, it is often inevitable that preferential flow will occur along the side of the column, since usually there are open spaces between the wall of the column and the soil material.

As with lysimeters, it is possible to monitor the leachate that drains from the column. It is also possible, however, to measure the concentrations at different depths in the column by placing sensors in the column.

17.7 A Closer look into Groundwater-Related Risk Assessment

17.7.1 Impact of Climate Change

A highly current environmental political issue is *climate change* caused by the increased emission of greenhouse gasses to the atmosphere. In addition to a significant worldwide increase in temperature, climate models are predicting serious changes in rainfall volume and distribution in the future. Although there is a lot of uncertainty in the model predictions, most climatic models point to an increase in frequency of intense precipitation episodes. As a consequence, dramatic changes in groundwater recharge and discharge regimes and in interactions between groundwater and surface water systems are expected as aquifers are replenished by effective rainfall. This phenomenon may impact the groundwater quality both in positive and in negative terms.

Another factor that influences groundwater flow and contaminant transport is *flooding*. Fluvial flooding occurs when rivers overflow and burst their banks, and sewage flooding occurs when sewage systems are not capable of evacuating the great volume of water during storms. As stated by the European Environment Agency, extreme floods are the most common natural disaster in Europe (Estrela et al. 2001) and probably in other parts of the world as well. Because of the increase of intense precipitation episodes, flooding is expected to happen more often in the future. It must be noted that extreme rainfall is indeed the main cause of flooding, but this factor is not the only one. The presence of vegetation and the soil capacity for water retention are other key factors with regard to flooding, and these factors are highly impacted by humans.

During flooding, the contaminants transported, mainly bound to particulate matter in the flood wave, can infiltrate and contaminate soil and, indirectly, groundwater. The properties of the vadouse zone are then essential in order to estimate the potential impact of flooding on the groundwater quality.

A special case of flooding happens when, due to occasional high rainfall events, the groundwater table rises above normal levels, and inundates roads and properties during weeks or even months (UK Groundwater Forum 2009). This phenomenon is known as *static groundwater flooding*, to differentiate it from transient groundwater flooding (US Geological Survey 2009) which happens when wet conditions persist for more than a year in terrains shaped from continental glaciers.

In some areas of the world, flood irrigation is practised. Especially in arid areas, groundwater is used for that purpose by means of the construction of irrigation wells at different depths. Such a process implies a disturbance of the natural flow regime and produces alterations in the recharge zone. As a consequence, a number of contaminants may be incorporated into the downward flow.

17.7.2 Mingling Groundwater Plumes

In many intensively used urban areas of the world, different groundwater plumes may intermingle in the subsurface. This phenomenon is important both in terms of groundwater as a protection target as well as in terms of groundwater as a means of transport (carrier) of contaminants with regard to transport to other protection targets. Intermingling of groundwater bodies is sometimes recognized but rarely incorporated in Groundwater-related Risk Assessment. There are several reasons for this. A political reason is that different groundwater plumes usually are related to another source and therefore to another ‘problem owner’, sometimes located in different jurisdictions. A technical reason is that the Risk Assessment with regard to a single groundwater plume is already a relatively unreliable process and assessing the risks from several mingling plumes is yet more difficult and, hence, may give rise to extreme uncertainties. Often the most reliable way to assess the risks from intermingling groundwater plumes, therefore, is by monitoring the contaminant transport pattern at strategic points or planes of compliance. However, after doing so, it is almost impossible to distinguish between the different sources of groundwater contamination and hence between the different polluters.

17.7.3 Risk Perception and Communication

Most of the issues concerning risk perception and risk communication described in Section 1.7.15 from a more general perspective also relate to groundwater contamination.

Analogous to the management of the upper soil layers, a *participatory approach* is needed for an effective implementation of water policies involving all relevant stakeholders. Communication and establishment of partnerships among stakeholders appear to be cornerstones of this participatory approach. One key objective in groundwater communication strategies should be focussed on providing environmental information, including uncertainty and risks, in an accessible way. Moreover, the trend towards regional approaches in groundwater Risk Management solutions requires a specific type of communication that includes negotiation and mediation (see Section 17.5.3).

Given the political and scientific interest in the impact of climate change on the hydrological cycle, communication about groundwater issues has been given a special dimension. In many developed countries the government provides educational materials in newspapers, magazines, and on television for the purpose of preparing the public for the consequences of climate change in terms of an increase in high precipitation events and flooding.

17.7.4 European Water Framework Directive and Groundwater Daughter Directive

On the 17th October 2006, an agreement was reached between the European Parliament and the Council of the European Commission for a new directive to protect groundwater from contamination (Quevauviller 2006). This new groundwater directive stems from the European Water Framework Directive of 2000 and sets Community Quality Standards for groundwater with respect to the major contaminants, nitrates and pesticides. It also creates the framework for setting any other Groundwater Quality Standards. The Groundwater Directive is based on a mixed regime, that is, compliance with Groundwater Quality Standards and measures to prevent (hazardous substances) or limit (non-hazardous substances) contaminant emissions to the groundwater. The European Member States are responsible for the establishment of some of these Groundwater Quality Standards with regard to local or regional conditions (e.g., the hydrogeological conditions, the nature of the topsoil, the interactions with the associated aquatic and terrestrial ecosystems, and the type of stress factors (Quevauviller (2006)).

In actual fact, the Groundwater Directive is a directive based on risks. Groundwater bodies must be classified by Member States ‘at risk of failing to achieve good chemical status’ or not. The investigation about whether the conditions for such good chemical status are met must include the risks from contaminants in the groundwater body as compared to those from potentially abstracted water for human consumption.

A site-specific Risk Assessment is needed to evaluate the risk posed by a certain activity and to ensure that inputs into groundwater do not cause deterioration. An additional trend assessment is also required to verify that an existing plume of contamination does not present a risk to human health or the environment.

17.8 Site-Specific Assessment of Exposure Through Contaminant Transport

In many countries, a tiered approach is used to be able to assess the site-specific human health risks due to contaminant transport in a scientifically based and efficient way. An example of this is given in Otte et al. (2007), who described the tiered approach used in the Netherlands within the framework of the determination of the urgency of remediation. In each tier successively the degree of conservatism decreases, while the site-specificity increases. As a consequence, the complexity, and hence the effort and finances needed, also increase in each tier. When in a specific tier unacceptable groundwater-related risks cannot be excluded, an assessment in the following tier then has to be performed. The underlying principle is: *simple when possible and complex when necessary.* The Dutch tiered approach is laid out as follows: Tier 0 relates to the ‘Initial Characterization’ which is based on three criteria: the presence of Non Aqueous Phase Liquids (NAPLs), vulnerable objects and/or a contaminated groundwater body in which the Intervention Value (the Dutch quality criterion for serious soil contamination) is exceeded in a groundwater volume of at least 6000 m^3. In Tier 1 a simple generic contaminant transport calculation is performed based on a conservative scenario. In Tier 2, a more complex site-specific calculation is performed based on the same concepts as in Tier 1, but instead of generic data, site-specific (i.e., measured) data needs to be used. Alternative options are also offered in Tier 2 to assess contaminant transport, for example, by extrapolating contaminant transport historical data to future transport estimates and including the possibility of degradation. Finally, in Tier 3, monitoring activities can be performed and complex numerical models can be used. Furthermore, leaching is taken into account in this tier, and special attention is given to Total Petroleum Hydrocarbons (TPH), a frequently occurring contaminant in groundwater.

References

Anderson J (2003) The environmental benefits of water recycling and reuse. Water Sci Technol: Water Supply 3(4):1–10

Baynes S (2004) Water efficient toilets: a canadian perspective. Canada Mortgage and Housing Corporation (electronic version). Retrieved 16 Oct 2009

Bruell CJ, Inyang HI (2000) The fundamentals of remediation in rock masses, chapter 1. In: Inyang HI, Bruell CJ (eds) Remediation in rock masses. ASCE, pp 1–11

Burke JJ, Moench MH (2000) Groundwater and society: resources, tensions and opportunities. United Nations Publication ST/ESA/205, New York

Celico F, Musilli I, Naclerio G (2004) The impacts of pasture- and manure-spreading on microbial groundwater quality in carbonate aquifers. Environ Geol 46(2):233–236

Chang CM, Wang MK, Chang TW, Lin C, Chen YR (2001) Transport modeling of copper and cadmium with linear and nonlinear retardation factors. Chemosphere 43(8): 1133–1139

Chisala BN, Lerner DN (2008) Sewage risks to urban groundwater. Science Report SC030134, Environment Agency, Bristol. 36 pp. ISBN 978-1-84432-820-8

Community Development Services Department, County of Humboldt (2009) http://co.humboldt.ca.us/planning. Retrieved 24 Dec 2009

Darcy HPG (1856) The public fountains in the city of Dijon (in French). Dalmont, Paris

Deeb RA, Chu K-H, Shih T, Linder S, Suffet I, Kavanaugh MC, Alvarez-Cohen L (2003) MTBE and other oxygenates: environmental sources, analysis, occurrence, and treatment. Environ Eng Sci 20(5):433–447

De la Orden Gómez JA (2006) The evaluation of the artificial recharge through wells with horizontal galleries in the aquifer Plana de Gandía-Denia, Municipalities of Vergel and Els Poblets (Alicante) (in Spanish). Internal Report of the Specific Agreement IGME-Diputación de Alicante

Environment Agency (2009) Underground, under threat, groundwater protection: policy and practice, part 1 – overview

Eurogeosurveys (2009) http://www.eurogeosurveys.org/download/Groundwater_Policy_bis.pdf. Retrieved 30 Dec 2009

European Commission (2008) Groundwater protection in Europe, the new groundwater directive – consolidating the EU regulatory framework, a publication of the European Communities

Estrela T, Menendez M, Dimas M (2001) Sustainable water use in Europe. Part 3. Extreme hydrological events: floods and droughts. Environmental issue report no. 21, European Environment Agency, Copenhagen, Denmark

Ezsoftech (2009) http://www.ezsoftech.com/hajj/hajj_article2.asp. Retrieved 30 Dec 2009

Field MS, Pinsky PF (2000) A two-region nonequilibrium model for solute transport in solution conduits in karstic aquifers. J Contamin Hydrol 44(3–4):329–351

Foster SSD, Chilton PJ (2003) Groundwater: the processes and global significance of aquifer degradation. Phil Trans R Soc Lond B 358:1957–1972. doi:10.1098/rstb.2003.1380

Grima J, Chacon E, Ballesteros B, Rodríguez R, Mejia JA (2009) Communication of groundwater realities based on assessment and monitoring data. Wiley, The Atrium, Southern Gate, Chichester, PO19 8SQ, England

Kale SP, Puetz T, Burauel P, Ophoff H, Fuehr F (2001) Long term fate of $_{14}$C-PCBs and $_{14}$C-PAHs in lysimeter soil. J Nucl Agric Biol 30(2):74–82

Kelsh MA, Buffler PA, Daaboul JJ, Rutherford GW, Lau EC, Barnard JC, Exuzides AK, Madl AK, Palmer LG, Lorey FW (2003) Primary congenital hypothyroidism, newborn thyroid function, and environmental perchlorate exposure among residents of a Southern California community. J Occup Environ Med 45(10):1116–1127

Kolb A, Püttmann W (2006) Comparison of MTBE concentrations in groundwater of urban and nonurban areas in Germany. Water Res 40(9):3551–3558

Leijnse A, Hassanizadeh SM (1994) Model definition and model validation. Adv Water Res 17(33):197–200

Lerner DN (2002) Identifying and quantifying urban recharge: a review. Hydrogeol J 10(1):143–152

Lerner DN, Tellam JH (1992) The protection of urban groundwater from pollution. J Inst Water Environ Manage 6(1):28–36

López-Geta JA, Fornés JM, Ramos G, Villarroya F (2006) Groundwater. A natural underground resource. Environmental education, IGME, UNESCO and Fundación Marcelino Botín, Spain, Madrid. ISBN: 84-7840-618-2

Lovanh N, Zhang Y-K, Heathcote RC, Alvarez PJJ (2000) Guidelines to determine site-specific parameters for modeling the fate and transport of monoaromatic hydrocarbons in groundwater. Report of the University of Iowa, October, 2000

MacQuarrie KTB, Mayer KU (2005) Reactive transport modeling in fractured rock: a state-of-the-science review. Earth-Sci Rev 72:189–227

Mandocdoc MC, Primo David C (2008) Dieldrin contamination of the groundwater in a former US military base (Clark air base, Philippines, CLEAN – SOIL). Air, Water 36(10–11): 870–874

Mayer KU, Frind EO, Blowes DW (2002) Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. Water Resour Res 38(9):1174

Mayes MA, Jardine PM, Larsen IL, Brooks SC, Fendorf SE (2000) Multispecies transport of metal–EDTA complexes and chromate through undisturbed columns of weathered fractured saprolite. J Contamin Hydrol 45(3–4):243–265

McCullough J, Hazen TC, Benson SM, Metting FB, Palmisano AC (1999) Bioremediation of metals and radionucleids. What it is and how it works. Office of biological and environmental research of the U.S. Department of Energy's Office of Science. NABIR primer LBNL-42595. Available at http://www.lbl.gov/NABIR/primer. Retrieved 30 Dec 2009

Meinrath G, May PM (2002) Thermodynamic prediction in the mine water environment. Mine Water Environ 21:24–35

Moran MJ, Lapham WW, Rowe BL, Zogorski JS (2007) Volatile organic compounds in ground water from rural private wells, 1986 to 1999. J Am Water Resou Assoc 40(5): 1141–1157

Optimum Population Trust (OPT) (2009) http://www.optimumpopulation.org/opt.more.water.html. Retrieved 9 Nov 2009

Otte PF, Zijp MC, Lijzen JPA, Swartjes FA, Verschoor AJ (2007) A tiered approach to assess risk due to contaminant migration in groundwater. RIVM report 711701056/2007. RIVM, Bilthoven, The Netherlands.

Quevauviller P (2006) Agreement on new EU groundwater directive. J Soils Sediments 6(4):254

Rahman MM, Chowdhury UK, Mukherjee SC, Mondal BK, Paul K, Lodh D, Biswas BK, Chanda CR, Basu GK, Saha KC, Roy S, Das R, Palit SK, Quamruzzaman Q, Chakrabort D (2001) Chronic arsenic toxicity in Bangladesh and West Bengal, India – a review and commentary. Clin Toxicol 39(7):683–700. doi:10.1081/CLT-100108509

Santos A, Alonso E, Callejón M, Jiménez JC (2002) Distribution of Zn, Cd, Pb and Cu metals in groundwater of the Guadiamar river basin. Water Air Soil Pollut 134(104):273–283

Schelwald-van der Kley AJM, Reijerker L (2009) Water: A way of life. Taylor and Francis, Leiden

Simmons CT, Fenstemaker TR, Sharp JM Jr (2001) Variable-density groundwater flow and solute transport in heterogeneous porous media: approaches, resolutions and future challenges. J Contamin Hydrol 52(1–4):245–275

Swartjes F, Sanders R, Tiktak A, Van der Linden T (1993) Modelling of leaching and accumulation of pesticides: module selection by sensitivity analysis. In: Del Re AAM, Capri E, Evans SP, Natali P, Trevisan M (eds.) Proceedings IX simposium pesticide chemistry: mobility and degradation of xenobiotics, pp 167–181

Tatalovich ME, Lee YKY, Chrysikopoulos VC (2000) Modeling the transport of contaminants originating from the dissolution of DNAPL pools in aquifers in the presence of dissolved humic substances. Transport Porous Media 38:93–115

Tomer MD, Burkart MR (2003) Long-term effects of nitrogen fertilizer use on ground water nitrate in two small watersheds. J Environ Qual 32:2158–2171

UK Groundwater Forum (2009) http://www.groundwateruk.org/Groundwater-drought-in-the-UK.aspx. Retrieved 30 Dec 2009

United Nations (1987) Report of the world commission on environment and development, A/RES/42/187, 96th plenary meeting, 11 December 1987

US EPA (2008) U.S. EPA's 2008 report on the environment (final report), U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-07/045F (NTIS PB2008-112484)

US Geological Survey (2009) http://pubs.usgs.gov/fs/2006/3026/. Retrieved 10 Sep 2009

Vrba J, Lipponen A (eds) (2007) Groundwater resource sustainability indicators. Groundwater indicators working group, UNESCO, IAEA, IHA, IHP – Viseries on groundwater No. 14.

Zektser IS, Everet LG (eds) (2004) Groundwater resources of the world and their use. UNESCO. IHP-VI Series on Groundwater No. 6

Zheng C, Gorelick SM (2005) Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale. Ground Water 41(2):142–155

Chapter 18
Leaching of Contaminants to Groundwater

Dirk Mallants, Martinus Th. Van Genuchten, Jiří Šimůnek, Diederik Jacques, and Suresh Seetharam

Abstract In this chapter the water flow and contaminant transport processes in the unsaturated or vadose zone are described. These processes include water retention and hydraulic conductivity, evapotranspiration, preferential flow, root water uptake (water flow) and diffusion, dispersion, advection and volatilization (contaminant transport). The equation governing transport of dissolved contaminants in the vadose zone is obtained by combining the contaminant mass balance with equations defining the total concentration of the contaminant and the contaminant flux density. Further attention is this chapter is given to nonequilibrium transport, stochastic models, multicomponent reactive solute transport, multiphase flow and transport. Mathematical models should be critical components of any effort to understand and predict site-specific subsurface water flow and contaminant transport processes. Generally, models range from relatively simple analytical approaches for analyzing contaminant transport problems during one-dimensional steady-state flow, to sophisticated numerical models for addressing multi-dimensional variably-saturated flow and contaminant transport problems at the field scale. An overview is given of several existing analytical and numerical models. Moreover, several applications to unsaturated flow and geochemical transport modeling are presented in this chapter.

Contents

D. Mallants (✉)
Belgian Nuclear Research Centre SCK•CEN, Mol, Belgium
e-mail: dmallant@sckcen.be

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_18,

18.1 Introduction

Many contaminants are finding their way into the subsurface environment, either intentionally applied, for example in agricultural operations, or unintentionally released from leaking industrial and municipal waste disposal sites, or from other sources (Šimůnek and Van Genuchten 2006). A broad range of fertilizers, pesticides and fumigants are now routinely applied to agricultural lands, thus making agriculture one of the most important sources for non-point source contamination. The same is true for salts and toxic trace elements, which are often an unintended consequence of irrigation in arid and semi-arid regions. Agriculture also increasingly uses a variety of pharmaceuticals and hormones in animal production many of which, along with pathogenic microorganisms, are being released to the environment through animal waste. Similar problems arise with point-source contamination from industrial and municipal waste disposal sites, leaking underground storage tanks, chemicals spills, nuclear waste repositories, and mine tailings, among other sources.

Mathematical models should be critical components of any effort to understand and predict site-specific subsurface water flow and contaminant transport processes. For example, models can be helpful tools for designing, testing and implementing soil, water and crop management practices that minimize soil and water contamination. Models are equally needed for designing or remediating industrial waste disposal sites and landfills, or for long-term stewardship of nuclear waste repositories. A large number of specialized numerical models now exist to simulate the different processes at various levels of approximation and for different applications.

Increasing attention is being paid recently to the *unsaturated or vadose zone* where much of the subsurface contamination originates, passes through, or can be eliminated before it contaminates groundwater, surface and subsurface water resources. Sources of contamination often can be more easily remediated in the vadose zone, before contaminants reach the underlying groundwater. The focus of this chapter thus will be on conceptual and mathematical descriptions of water flow and especially transport processes in the predominantly unsaturated or variably-saturated vadose zone. The vadose zone is defined here as the zone between the land surface and the permanent (seasonal) groundwater table. The vadose zone is usually only partially saturated, although saturated regions may exist, such as perched water above a low-permeable fine-textured (clay) layer or a saturated zone behind the infiltration front during or after a high-intensity rainfall event.

Since the transport of contaminants is closely linked with the water flux in soils and rocks making up the vadose zone, any quantitative analysis of contaminant transport must first evaluate water fluxes into and through the vadose zone. Water typically enters the vadose zone in the form of rainfall or irrigation (Fig. 18.1), or by means of industrial and municipal spills. Some of the rainfall or irrigation water may be intercepted on the leaves of vegetation. If the rainfall or irrigation intensity is larger than the infiltration capacity of the soil, water will be removed by surface runoff, or will accumulate at the soil surface until it evaporates back to the atmosphere or infiltrates into the soil. Some of the water that infiltrates into the soil profile may be taken up by plant roots and eventually returned to the atmosphere by plant transpiration. The processes of evaporation and transpiration are often combined into the single process of evapotranspiration. Only water that is not returned to the atmosphere by evapotranspiration may percolate to the deeper vadose zone and eventually reach the groundwater table. If the water table is close enough to the soil surface, the process of capillary rise may move water from the groundwater table through the capillary fringe towards the root zone and the soil surface.

Because of the close linkage between water flow and contaminant transport, we will first briefly focus on the physics and mathematical description of water flow in the vadose zone (Section 18.2). An overview is given of the governing equations for water flow, while a comprehensive example is used to illustrate water content and pressure head distributions in single- and multi-layered soil profiles following steady-state infiltration (Section 18.2.8). This is followed by a discussion of the equations governing contaminant transport (Section 18.3) where we review the standard equilibrium formulations for contaminant transport (Section 18.3.2) as well as alternative non-equilibrium models (Section 18.3.3). We also briefly discuss possible formulations for colloid-facilitated transport (Section 18.3.3.3), stochastic approaches for contaminant transport (Section 18.3.4), and multicomponent geochemical transport (Section 18.3.5). This is followed by a discussion of analytical (Section 18.4) and numerical (Section 18.5) approaches for solving the governing flow and/or transport equations, and an overview of computer models currently

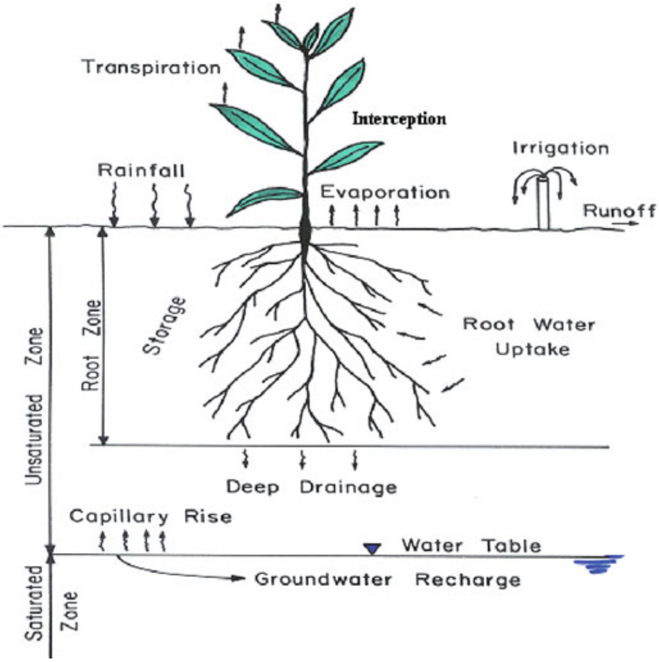


Fig. 18.1 Schematic of water fluxes and various hydrologic components in the vadose zone (Šimůnek and Van Genuchten 2006)

available for simulating vadose zone flow and transport processes (Sections 18.4.2 and 18.5.2). Several sections of this text are adapted from a recent book chapter by Šimůnek and Van Genuchten (2006).

18.2 Variably Saturated Water Flow

In this Section we briefly present the equations governing unsaturated water flow in the subsurface. Traditionally, descriptions of variably saturated flow in soils are based on the Richards (1931) equation, which combines the Darcy-Buckingham equation for the fluid flux with a mass balance equation. The Richards equation typically predicts a uniform flow process in the vadose zone, although possibly modified macroscopically by assuming the presence of spatially variable soil hydraulic properties (e.g., as dictated by the presence of different soil horizons, but possibly also varying laterally). Unfortunately, the vadose zone can be extremely heterogeneous at a range of scales, from the microscopic (e.g., pore scale) to the macroscopic (e.g., field or larger scale). Some of these heterogeneities can lead to a preferential flow process that macroscopically is very difficult to capture with the standard Richards equation. One obvious example of preferential flow is the rapid movement of water and dissolved contaminants through macropores (e.g., between soil aggregates, or created by earthworms or decayed root channels, see e.g. Mallants et al. (1996a)) or rock fractures, with much of the water bypassing (short-circuiting) the soil or rock matrix. However, many other causes of preferential flow exist, such as flow instabilities caused by soil textural changes or water repellency (Hendrickx and Flury 2001; Ritsema and Dekker 2000; Šimůnek et al. 2003), and lateral funneling of water due to inclined or other textural boundaries (e.g., Kung 1990). Alternative ways of modeling preferential flow are discussed in a later section. Here we first focus on the traditional approach for uniform flow as described with the Richards equation.

18.2.1 Water Retention and Hydraulic Conductivity

18.2.1.1 Water Retention

Above the groundwater table, a zone of a few to several tens of meters occurs where part of the pore space is occupied by the air phase. This is the unsaturated or vadose zone, of which the upper part typically contains a soil profile. When the soil becomes drier owing to internal drainage and/or evapotranspiration, air replaces water first in the coarse parts of the pore space and at successively lower (negative) values of the water potential (see further) also in the finer pores. In the unsaturated zone, water is held in the soil pores by so-called surface-tension forces. In other words, capillary forces (and to a lesser extent also adsorption) bind water to solids. This leads to the existence of a negative pressure, also referred to as the suction, tension, or matric head (by definition pressures less than atmospheric are considered negative).

Capillary forces are the result of a complex set of interactions between the solid (particles) and liquid (pore water) phases involving the surface tension of the liquid phase, the contact angle between the solid phase and the liquid phase, and the diameter of pores. As a consequence of these forces, water will rise to a height H [L] when a capillary tube of radius R is placed into a water reservoir open to the atmosphere (Fig. 18.2). This capillary rise is given by the *Laplace equation*:

$$H = \frac{2\sigma \cos \gamma}{\rho_w g R} \tag{18.1}$$

where σ is the surface tension [MT^{-2}] (7.27×10^{-2} kg/s^2 at 20°C), γ is the contact angle, ρ_w is the density of the liquid phase [ML^{-3}] (998 kg/m^3 at 20°C), and g is the gravitational acceleration [LT^{-2}] (9.81 m/s^2). For water at 20°C with $\gamma = 0$, Eq. (18.1) can be simplified to $H \cong 1.5\times10^{-5}/R$ (with both H and R in m).

Since a soil can be viewed as a complex system containing pores of various diameters, water in those pores will rise to different heights (Fig. 18.3), and hence will be held with different potential energies. Because each soil has a different distribution of pore sizes, the distribution of water above the water reservoir will also be different. This simple conceptual model assumes that soil pores can be represented

Fig. 18.2 Capillary rise *H* of water in a cylindrical tube of radius *R* with a contact angle γ

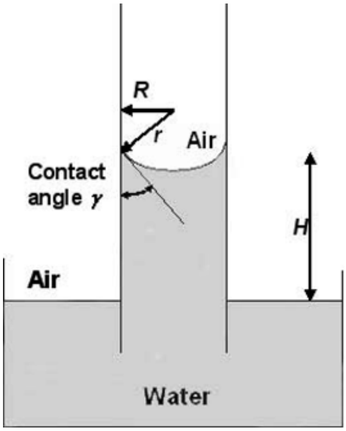
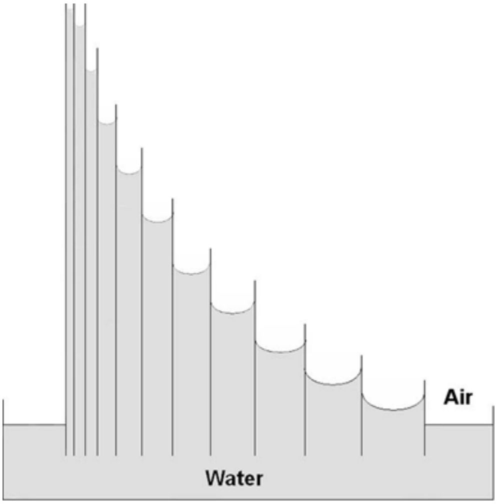


Fig. 18.3 Capillary rise of water in a set of cylindrical tubes of various radii



by an equivalent bundle of capillaries with identical water retention properties as the real soil. Such an approach with capillary tubes and a water reservoir can be used to evaluate the water content distribution in a soil above the groundwater table at equilibrium. The experimental curve that describes this relationship between the water content versus the height above the water table is called here the *water retention curve*. Many other names may be found in the literature, including pore water characteristic curve, capillary pressure-saturation relationship, and pF curve. The retention curve historically was often given in terms of pF, which is defined as the negative logarithm of the absolute value of the pressure head in cm. The water retention curve provides information on how tightly water is held in pores and how much work would need to be exerted to extract it from the different pores. The water retention curve thus characterizes the energy status of water in the soil, and is one of the two soil hydraulic functions necessary to describe the status and movement of water in the vadose zone. Since capillary rise depends on the radius of a particular capillary, the retention curve may also be interpreted as a curve that characterizes the distribution of pores of different radii in the soil.

The dependency of the fraction of water-filled porosity (i.e., the water content θ) on the soil matric potential ψ_m is formalized in the soil water retention curve, $\theta(\psi_m)$. This relation is of fundamental importance for the hydraulic characterization of a soil since it relates an energy density (potential) to a capacity quantity (water content). One may view the water retention curve as the curve that results when water is slowly removed from an initially water saturated soil until the soil becomes air saturated. This is the desorption curve. The adsorption (or absorption) curve describes the reverse process. Note that air can only enter the porous medium after the matric potential has fallen below a certain value ψ_{m0}, the so-called air-entry value. This value is determined by the largest pore of the porous medium open to air. Once air has entered the porous medium, the water content decreases monotonically with increasingly negative matric potentials ψ_m.

Rather than using the matric potential in the parameterization of the water characteristic, the pressure head h is often used instead:

$$h = -\frac{\Psi_m}{\rho_w g} \tag{18.2}$$

where h has SI units of meter (m) and dimensions of L; h represents the energy state of pore water and is expressed as energy per unit weight.

Figure 18.4 shows water retention curves for three different textural classes, i.e., for a clay, a loam, and a sand, as given by Carsel and Parish (1988) using the soil hydraulic parameters given in Table 18.1 (to be discussed later). Notice that the sand loses its water relatively quickly (at small negative pressure heads) and abruptly above the water table (the pressure head $h = 0$ cm at the water table), while the more fine-textured loam and especially the clay soil lose their water much more gradually. This reflects the pore-size distribution of a particular soil textural class. While the majority of pores in coarse-textured soils (such as sand and gravel) have relatively large diameters and thus drain at relatively small negative pressures, the majority of pores in fine-textured soils (such as clays, silty clays and clay loams) do not drain until very large tensions (negative pressures) are applied.

Commonly used mathematical expressions for the retention curve, $\theta(h)$, are the Van Genuchten (Van Genuchten 1980a) and Brooks and Corey (1964) equations since they permit a relatively good description of $\theta(h)$ for many soils using only a limited number of parameters. The Van Genuchten soil moisture retention characteristic is defined as:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{(1 + |ah|^n)^m} \tag{18.3}$$

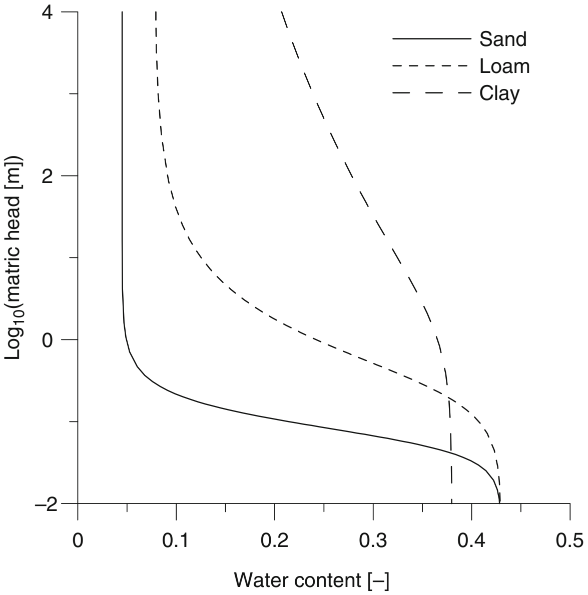


Fig. 18.4 Example of retention curves for sand, loam and clay textural classes (based on Carsel and Parish (1988) soil hydraulic parameters, Table 18.1)

Table 18.1 Soil hydraulic parameters for the analytical functions of Van Genuchten (1980a) for the twelve textural classes of the USDA soil textural triangle (Carsel and Parish 1988)

Textural class	θ_r [L^{3}L^{-3}]	θ_s [L^{3}L^{-3}]	α [cm^{-1}]	n [–]	K_s [cm d^{-1}]
Sand	0.045	0.430	0.145	2.68	712.8
Loamy Sand	0.057	0.410	0.124	2.28	350.2
Sandy Loam	0.065	0.410	0.075	1.89	106.1
Loam	0.078	0.430	0.036	1.56	24.96
Silt	0.034	0.460	0.016	1.37	6.00
Silty Loam	0.067	0.450	0.020	1.41	10.80
Sandy Clay Loam	0.100	0.390	0.059	1.48	31.44
Clay Loam	0.095	0.410	0.019	1.31	6.24
Silty Clay Loam	0.089	0.430	0.010	1.23	1.68
Sandy Clay	0.100	0.380	0.027	1.23	2.88
Silty Clay	0.070	0.360	0.005	1.09	0.48
Clay	0.068	0.380	0.008	1.09	4.80

where θ_r is the residual water content [L^{3}L^{-3}], θ_s is the saturated water content [L^{3}L^{-3}], and α [L^{-1}], n [−] and m (= 1−1/n) [–] are shape parameters. The dependency of the Van Genuchten model on the parameters α and n is shown in Fig. 18.5. The curves in this figure are plotted in terms of effective saturation given by S_e=(θ–θ_r)/(θ_s−θ_r).

As a first approximation and on intuitive grounds, $\theta_s = \eta$ (total porosity) and $\theta_r = 0$. In reality, however, the saturated water content θ_s of soils is often smaller than the total porosity η because of entrapped and dissolved air. The residual water content θ_r is generally larger than zero, because of the presence of adsorbed water.

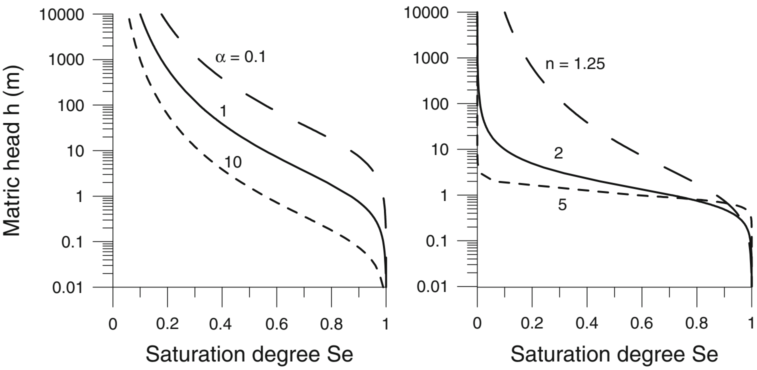


Fig. 18.5 Dependency of soil water characteristic curves on Van Genuchten parameters α (keeping n fixed at 1.25) (*left*) and n (keeping α fixed at 1) (*right*). Saturation degree S_e is used rather than water content

While some definitions of θ_r have been proposed (e.g., the water content where the hydraulic conductivity becomes first zero during drying), θ_r in practice is best treated as a fitting parameter without any physical significance.

Another parametric description of the moisture retention characteristic is the equation of Brooks and Corey (1964):

$$\theta(h) = \begin{cases} \theta_r + (\theta_s - \theta_r)\left|\frac{h_e}{h}\right|^n & h < h_e \\ \theta_s & h \geq h_e \end{cases} \tag{18.4}$$

where h_e is the air-entry value [L] and n is the pore-size distribution index that characterizes the width of the pore-size distribution (or the steepness of the retention function). Like n in Eq. (18.3), the value of λ is relatively large for soils with a relatively uniform pore-size distribution, but small for soils having a wide range of pore sizes.

A large number of experimental techniques can be used to measure the water retention curves. These techniques include methods using a hanging water column, pressure cells, pressure plate extractors, suction tables, soil freezing, and many other approaches (Dane and Hopmans 2002). Alternatively, soil hydraulic parameters may be obtained with pedotransfer functions using information on textural class or textural fractions (and some additional information). Pedotransfer functions are statistical relationships that use easily available soil data to estimate the soil hydraulic parameters (Bouma 1989; Leij et al. 2002; Pachepsky and Rawls 2004; Pachepsky et al. 2006).

18.2.1.2 Hydraulic Conductivity

Another important soil hydraulic property is the *unsaturated hydraulic conductivity function*. The *hydraulic conductivity* characterizes the ability of a soil to transmit water, and as such is inversely related to the resistance to water flow. The hydraulic conductivity depends on many factors, including the pore-size distribution of the porous medium, and the tortuosity, shape, roughness, and degree of interconnectedness of pores. The hydraulic conductivity decreases considerably as soil becomes unsaturated and less pore space is filled with water. The unsaturated hydraulic conductivity function gives the dependency of the hydraulic conductivity on the water content, $K(\theta)$, or pressure head, $K(h)$. The conceptual model that views the soil as a bundle of capillaries of different radii, as used earlier to explain water retention properties (Fig. 18.3), may be used also to evaluate the hydraulic conductivity function. The ability of each capillary tube to conduct water can be calculated using Poiseuille's flow equation. By adding the conductivity of all capillaries that are filled with water at a particular water content or pressure head, one obtains the hydraulic conductivity of the complete set of capillaries, and consequently of the soil itself.

Figure 18.6 presents examples of hydraulic conductivity functions for the sand, loam, and clay textural classes, again using the Carsel and Parish (1988) parameter values as listed in Table 18.1. The hydraulic conductivity curves in Fig. 18.6 are presented as functions of both the pressure head (left) and water content (right). Notice that the hydraulic conductivity at saturation is significantly larger for coarse-textured soils (sands and gravels) than for loams and clays. This difference is often several orders of magnitude. Also notice that the hydraulic conductivity may decrease several orders of magnitude as the soil becomes unsaturated. This decrease, when expressed as a function of the pressure head (Fig. 18.6, left), is much more significant for sands than for loams, and even more so than for clays. The decrease for coarse-textured soils (sands and gravel) is so large that at a certain pressure head the hydraulic conductivity becomes smaller than for heavy-textured soils (clays). These properties of the hydraulic conductivity function are often used in the design of

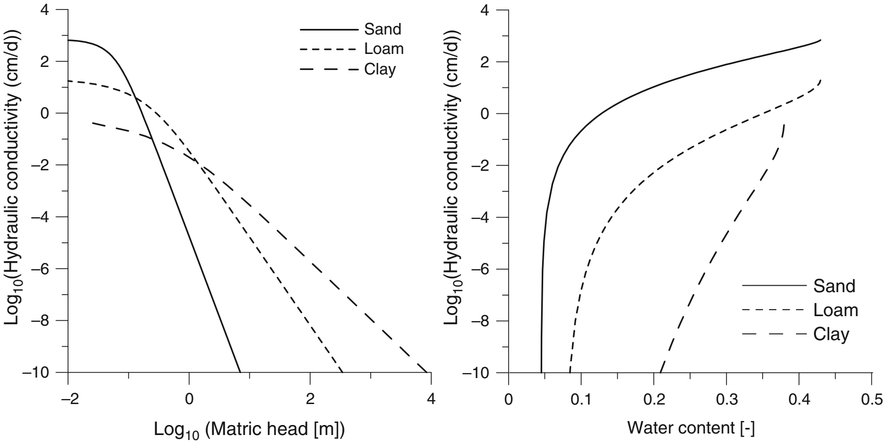


Fig. 18.6 Example of hydraulic conductivity curves for the sand, loam and clay soil textural classes (based on Carsel and Parish (1988) soil hydraulic parameters, Table 18.1)

engineered structures, such as capillary barriers (finer-textured soils above coarser-textured soils) in landfill covers to divert water from flowing through the underlying waste (e.g., Mallants et al. 1999), or having narrow sand or gravel layers at the soil surface to prevent or limit evaporation.

Similarly as for the water retention curve, analytical models are often used also for the hydraulic conductivity function. Many of these functions were derived using the pore-size distribution models of Burdine (1953) or Mualem (1976) in combination with one of the retention functions given earlier. The Brooks and Corey (1964) retention (Eq. (18.4)) is commonly associated with Burdine's pore-size distribution model, leading to the hydraulic conductivity function:

$$K = K_s S_e^{2/\lambda + l + 1} \tag{18.5}$$

in which K_s is the saturated hydraulic conductivity [LT^{-1}], $S_e{=}(\theta{-}\theta_r)/(\theta_s{-}\theta_r)$ is effective saturation, and l a pore-connectivity parameter assumed to be 2.0 in the original study of Brooks and Corey (1964).

The Van Genuchten (1980a) retention function (Eq. (18.3)) is similarly coupled mostly with the model of Mualem (1976) to give

$$K(h) = K_s S_e^l [1 - (1 - S_e^{1/m})^m]^2 \tag{18.6}$$

where

$$m = 1 - 1/n,\ n > 1 \tag{18.7}$$

The pore-connectivity parameter l in (Eq. (18.7)) was estimated by Mualem (1976) to be about 0.5 as an average for many soils. However, more recently, Schaap and Leij (2000) recommended using l equal to –1 as an appropriate value of most soil textures.

The saturated hydraulic conductivity K_s can be measured in the laboratory using either the constant or falling head method, and in the field using single or double ring infiltrometers, a constant head well permeameter, the auger-hole method, or the piezometer method (Dane and Topp 2002). Many laboratory and field methods are also available for measuring unsaturated hydraulic conductivities (Dane and Topp 2002).

18.2.2 Mass Balance Equation

Water flow in variably saturated rigid porous media (such as soils) is usually formulated in terms of a mass balance equation of the form

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - S \tag{18.8}$$

where θ is the volumetric water content [L^3L^{-3}], t is time [T], z is the spatial coordinate [L], q is the volumetric flux [LT^{-1}], and S is a general sink/source term [$L^3L^{-3}T^{-1}$], for example to account for root water uptake (transpiration).

Equation (18.8) is often referred to also as the *mass conservation equation* or the continuity equation. The mass balance equation in general states that the change in the water content (storage) in a given volume is due to spatial changes in the water flux (i.e., fluxes in and out of some small volume of soil) and possible sinks or sources within that volume. The mass balance equation must be combined with one or several equations describing the volumetric flux (q) to produce the governing equation for variably saturated flow.

For a soil that can be saturated or unsaturated, the flux is given by the equation:

$$q = -k(h)\frac{\partial h}{\partial z} + K(h) \tag{18.9}$$

where K is the unsaturated hydraulic conductivity [LT^{-1}], and h the pressure head. Eq. (18.9) is often referred to as the Darcy-Buckingham equation. The hydraulic conductivity in this equation is a function of the pressure head, h. In saturated systems, the conductivity becomes independent of the pressure head, in which case the equation reduces to Darcy law as:

$$q = -K_s\frac{\partial h}{\partial z} + K_s \tag{18.10}$$

where K_s is the saturated hydraulic conductivity [LT^{-1}]. The Darcy-Buckingham equation hence is formally similar to Darcy's equation, except that the proportionality constant (i.e., the unsaturated hydraulic conductivity) in the Darcy-Buckingham equation is a nonlinear function of the pressure head (or water content), while $K(h)$ in Darcy's equation is a constant equal to the saturated hydraulic conductivity, K_s (e.g., see discussion by Narasimhan 2005).

Combining the mass balance (Eq. (18.8)) with Darcy-Buckingham's law (Eq. (18.9)) leads to

$$\frac{\partial \theta(h)}{\partial t} = \frac{\partial}{\partial z}\left[K(h)\frac{\partial h}{\partial h} + K(h)\right] - S(h) \tag{18.11}$$

which was first formulated by Richards (1931) and is now generally referred to as the Richards equation. This partial differential equation is the equation governing water flow in the variably-saturated vadose zone. Because of its strongly nonlinear makeup, only a relatively few simplified analytical solutions can be derived. Most practical applications of Eq. (18.11) require a numerical solution, which can be obtained using a variety of methods such as finite differences or finite elements. Equation (18.11) is generally referred to as the mixed form of the Richards equation since it contains two dependent variables, i.e., the water content and the pressure head. Various other formulations of the Richards equation are possible.

18.2.3 Preferential Flow

Increasing evidence exists that water in many field soils does not move according to the uniform flow pattern typically predicted with the Richards equations (Flury et al. 1994; Hendrickx and Flury 2001). This is due to the presence of macropores, fractures or other structural voids or biological channels through which water and contaminants may move preferentially, while bypassing a large part of the matrix pore-space, called *preferential flow*. Preferential flow and transport processes are most frustrating in terms of hampering accurate predictions of contaminant transport in soils and fractured rocks (Šimůnek and Van Genuchten 2006). Contrary to uniform flow, preferential flow results in irregular wetting of the soil profile as a direct consequence of water moving faster in certain parts of the soil profile than in others. Hendrickx and Flury (2001) defined preferential flow as constituting all phenomena where water and contaminants move along certain pathways, while bypassing a fraction of the porous matrix. Water and contaminants for these reasons can move quickly to far greater depths, and much faster, than would be predicted with the Richards equation describing uniform flow.

The most important causes of preferential flow are the presence of macropores and other structural features, development of flow instabilities (i.e., fingering) caused by profile heterogeneities or water repellency (Hendrickx et al. 1993), and funneling of flow due to the presence of sloping soil layers that redirect downward water flow. While the latter two processes (i.e., flow instability and funneling) are usually caused by textural differences and other factors at scales significantly larger than the pore scale, macropore flow and transport are usually generated at the pore or slightly larger scales, including scales where soil structure first manifests itself, mostly the pedon scale (Šimůnek et al. 2003).

Uniform flow in granular soils and preferential flow in structured media (both macroporous soils and fractured rocks) can be described using a variety of single-porosity, dual-porosity, dual-permeability, multi-porosity, and/or multi-permeability models (Gerke and Van Genuchten 1993; Gwo et al. 1995; Jarvis et al. 1998; Mallants et al. 1996b, 1996c; Pruess and Wang 1987; Richards 1931; Šimůnek et al. 2003, 2005). While single-porosity models assume that a single pore system exists that is fully accessible to both water and contaminant, dual-porosity and dual-permeability models assume that the porous medium consists of two interacting pore regions, one associated with the inter-aggregate, macropore, or fracture system, and one comprising the micropores (or intra-aggregate pores) inside soil aggregates or the rock matrix. Whereas dual-porosity models assume that water in the soil or rock matrix is stagnant, dual-permeability models allow also for water flow within the matrix.

18.2.4 The Evapotranspiration Process

Evapotranspiration (*ET*) is the process by which water is transported from the earth surface (i.e., the plant-soil system) to the atmosphere by evaporation (*E*) from surfaces (soils and wet vegetation) and by transpiration (T) from plants through stomata in the plant leaves (Fig. 18.1). Evaporation is the physical process by which water is transformed into water vapour and is removed from the evaporating surface by mass transfer. Transpiration is the process by which water in plant tissues is transformed into water vapour and removed towards the atmosphere. Important weather parameters affecting ET are radiation, air temperature, humidity and wind speed (see further). Another important factor is the amount of water available at the evaporating surface (e.g., the soil surface), or in the soil for uptake by plant roots. The available water is determined by such factors as soil type (texture), depth to ground water, irrigation and management practices. Root water uptake is negatively influenced by water logging (i.e., when the soil is completely or almost completely saturated), and pore water salinity. Crop characteristics such as crop type, variety, development stage, crop height, crop roughness, reflection, ground cover and crop rooting characteristics all influence the resistance to transpiration, and hence ET.

To remove the effects due to soil type, management and crop factors on calculations of the evaporative demand of the atmosphere, the evapotranspiration rate is generally calculated for a reference surface not short of water. This ET is called the reference evapotranspiration, ET_o, and usually calculated following guidelines of the FAO56 paper by Allen et al. (1998) using the Penman-Monteith equation (Monteith 1965) (see Section 18.2.6). The reference surface is a hypothetical reference crop (grass) with an assumed height of 0.12 m, a fixed surface resistance of 70 s m^{-1} and an albedo of 0.23. ET_o thus only depends on climatic parameters and can be considered as a climatic parameter expressing the evaporating power of the atmosphere at a specific location and time of the year (Allen et al. 1998).

Since actual ground cover, canopy properties and aerodynamic resistance of the crop are different from those used for calculating ET_o, the evapotranspiration rate under standard conditions (i.e., of a large field under excellent agronomic and pore water conditions) for a specific crop, ET_c, is required for specific applications. ET_c can be obtained either by using specific crop parameters (e.g., albedo, aerodynamic and canopy surface resistances) in the Penman-Monteith equation, or by multiplying ET_o with a crop coefficient K_c. The crop coefficient incorporates four primary characteristics that distinguish a specific crop from the reference crop (crop height, albedo, canopy resistance, and evaporation from the soil) (Allen et al. 1998) and is determined by crop type, climate, soil evaporation and crop growth stages. Consequently, K_c coefficients change during a growing season.

Actual evapotranspiration under field conditions, ET_a, takes into account non-ideal (non-standard) conditions such as water or salinity stress. Under dry soil conditions, water flow in a soil can be too slow to satisfy the evaporative demand. Similarly as for transpiration, very dry or very wet soil conditions, or high salt concentrations, impose water and salinity stresses and reduce root water uptake. Allen et al. (1998) related ET_a to ET_o by means of a water stress coefficient and/or by adjusting K_c for all kinds of stresses. In this chapter, a mechanistic approach to determine ET_a is described using water flow in the soil and water uptake by plant roots, the latter incorporating empirical water stress reduction functions (see Section 18.2.5).

18.2.5 Penman-Monteith Equation for Evapotranspiration

For evapotranspiration to occur, three conditions are needed in the soil-plant-atmosphere system (Jensen et al. 1990):

1. a supply of water must be available;
2. energy must be available to convert liquid water into vapour water;
3. a vapour pressure gradient must exist to create a flux from the evaporating surface to the atmosphere.

Penman (1948) proposed a combination method by introducing an energy balance (condition 2) and a mass transfer term in an aerodynamic formula (condition 3) into a single equation to calculate *ET*. Penman's method was developed to calculate *E* as open water evaporation. Written as the weighted sum of the rates of evaporation due to net radiation, E_r (MJ m^{-2} d^{-1}), and turbulent mass transfer, E_a (MJ m^{-2} d^{-1}), Penman's equation for the evaporative latent heat flux, λE (MJ m^{-2} d^{-1}), is:

$$\lambda E = \frac{\Delta}{\Delta + \gamma} E_r + \frac{\gamma}{\Delta + \gamma} E_a \tag{18.12}$$

where λ is the latent heat of vaporization (MJ kg^{-1}), Δ is the slope of the vapour pressure curve (kPa °C^{-1}) and γ is the psychrometric constant (kPa °C^{-1}). E_r is given by:

$$E_r = R_n - G \tag{18.13}$$

where R_n is the net radiation flux (MJ m^{-2} d^{-1}) and G is the sensible heat flux into the soil (MJ m^{-2} d^{-1}), and E_a by:

$$E_a = W_f(e_s - e_a) \tag{18.14}$$

where e_s and e_a are the saturation and actual vapour pressures, respectively (kPa), (e_s–e_a) is the saturation vapour pressure deficit, and W_f is a wind function (MJ d^{-1} kPa^{-1}). A linear wind function was found to be adequate, defined as (Allen 2001):

$$W_f = K_w(a_w + b_w u_2) \tag{18.15}$$

where K_w is a units conversion factor [6.43 for ET_0 in mm d^{-1}], and a_w and b_w are empirical wind function coefficients often obtained by regional or local calibration.

The Penman method to estimate the evaporation from open water is then:

$$E = \frac{1}{\lambda}\left(\frac{\Delta}{\Delta + \gamma}(R_n - G) + \frac{\gamma}{(\Delta + \gamma)} K_w(a_w + b_w u_2)(e_s - e_a)\right) \tag{18.16}$$

Penman (1948) derived this equation for open water evaporation. Evaporation from bare soil, wet soil and grasses is obtained as a fraction of E. Bulk surface resistances from the soil and crop is not explicitly accounted for, but are incorporated in the wind function. Resistance factors are incorporated in Penman-based equations to include the resistance of vapour flow through stomata openings, total leaf area and the soil surface (the surface resistance, r_s), and the resistance from the vegetation upwards involving friction from air flowing over vegetative surfaces (aerodynamic resistance, r_a) (Allen et al. 1998). The Penman-Monteith equation (Monteith 1965) for evaporation from bare soil, wet soil and grasses, ET, is given in the ASCE standard form (Allen et al. 1998) as:

$$ET = \frac{1}{\lambda}\left(\frac{\Delta(R_n - G) + K_{time}\rho_a c_p \frac{(e_s - e_a)}{r_a}}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)}\right) \tag{18.17}$$

where ρ_a is the mean air density at constant pressure (kg m^{-3}), c_p is the specific heat of air (MJ kg^{-1} °C^{-1}), r_a and r_s are the aerodynamic and (bulk) surface resistances, respectively (s m^{-1}), and K_{time} is a units conversion factor (86400 s d^{-1} when ET is expressed in mm d^{-1}). For a more detailed definition and discussion of Eq. (18.17) and its parameters the reader is referred to Allen et al. (1998).

18.2.6 FAO56 Reference Evapotranspiration

Allen et al. (1998) calculated the reference evapotranspiration, ET_0, using the ASCE Penman-Monteith equation (Eq. (18.17)) for a hypothetical reference surface or reference crop defined as “a cropped soil with an assumed crop height of 0.12 m, a fixed surface resistance of 70 s m^{-1} and an albedo of 0.23”, with climatological parameters measured at a reference level of 2 m above the soil surface. The popularly used FAO56 Penman-Monteith equation is defined as (Allen et al. 1998):

$$ET_0 = \frac{0.408\Delta\,(R_n - G) + \gamma\frac{900}{T + 273}u_2\,(e_s - e_a)}{\Delta + \gamma\,(1 + 0.34u_2)} \tag{18.18}$$

To calculate ET_0 using the FAO56 Penman-Monteith equation on a daily basis, such as implemented for instance in the ET-REF program (Allen 2001), site-specific data (altitude above sea level and latitude) and climatological data (temperature, humidity, radiation, and wind speed) are required. The altitude above sea level determines the local average value of atmospheric pressure. The latitude is needed to compute extraterrestrial radiation.

The potential evapotranspiration of a particular crop or vegetation, ET_c, is then obtained by multiplying ET_0 with a crop coefficient, K_c. ET_c is divided between evaporation of the intercepted water ($E_{p,I}$), potential soil evaporation ($E_{p,s}$) and potential transpiration (T_p) using procedures outlined below.

Daily values of the actual soil evaporation rate, $E_{a,t}$ [L], and actual transpiration rate, $T_{a,t}$ [L], may be calculated using HYDRUS-1D (Šimůnek et al. 2005). The input variables for HYDRUS-1D are time series of daily values for the throughfall (e.g., precipitation or irrigation reaching the soil surface), T, potential soil evaporation, $E_{p,s}$, and potential transpiration, T_p.

18.2.7 Root Water Uptake

The sink term S in Eq. (18.8) is defined as the volume of water extracted from a unit volume of soil per unit time by the roots. The potential *root water uptake rate* S_p(z) is often obtained by multiplying a normalized water uptake distribution $b(z)$ [L^{-1}] with the potential transpiration rate T_p [LT^{-1}] as follows:

$$S_p(z) = b(z)T_p \tag{18.19}$$

The function $b(z)$ may be obtained from the root distribution with depth:

$$b(z) = \frac{b'(z)}{\int\limits_{L_R} b'(z)} \tag{18.20}$$

where $b'(z)$ is the root distribution function and L_R is the soil root zone. Note that $b'(z)$ can be of any form. T_p depends on climate conditions and vegetation (leaf area index, crop coefficients – see further). The actual root water uptake rate S(z) may be obtained by multiplying $S_p(z)$ with a root water stress response function (e.g., Feddes et al. 1978; Van Genuchten 1987) to account for a possible reduction in root water uptake due to water stress conditions in the soil profile:

$$S(h,z) = \alpha(h)S_p = \alpha(h)b(z)T_p \tag{18.21}$$

where $\alpha(h)$ is the water stress response function as a function of the pressure head.

To obtain the actual transpiration rate T_a of the vegetation, the actual root water uptake $S(h,$z) (Eq. (18.21)) is integrated over the rooting depth:

$$T_a = T_p \int\limits_{L_R} \alpha(h,z)b(z)dz \tag{18.22}$$

18.2.8 Application: Numerical Simulations of Variably Saturated Flow in a Soil Profile

Analysis of water flow in soils can be based on experimental exploration and/or numerical simulation. Experimental studies yield valuable insight into the physical processes, but their implementation is often complicated, time consuming, and costly. As an alternative, or a useful complement, numerical simulations of the water regime can provide necessary information to analyse flow phenomena at a very high spatio-temporal resolution and at reasonable cost. Also, simulations may be used to plan and interpret experiments and to predict future flow behaviour for complex conditions.

In the following examples, numerically simulated water flow will be discussed for different conditions of increasing complexity. We start the discussion with one-dimensional flow in a hypothetical single-layer (homogeneous) soil profile, and then move to a two-layered soil. Note that we will not calculate water uptake by roots explicitly, but rather apply a net rainfall rate, which is equal to the total rainfall rate minus the actual evapotranspiration (runoff is assumed zero here, but could be accounted for if needed).

The simulations for the hypothetical soil made use of the Van Genuchten-Mualem hydraulic functions without hysteresis. Figure 18.7 shows the soil water retention and hydraulic conductivity functions. All simulations were carried out with the HYDRUS-1D finite element code (Šimůnek et al. 2005). As top boundary condition for the hypothetical soil profile we implemented a constant rainfall rate of 2×10^{-6} m/s (or 7.2×10^{-3} m/h). This value corresponds to one tenth of the saturated hydraulic conductivity of the sand material. As bottom boundary we assumed a groundwater table 4 m below surface. This condition is mathematically implemented by fixing the pressure head equal to zero at the water table. As initial condition for the pressure head we used $h = -z - 4$ m, where z is the depth below surface (the vertical axis is positive upward, with the reference depth $z = 0$ at the soil surface, see Fig. 18.8). At the soil surface the pressure head initially is hence $0 - 4 = -4$ m, with h decreasing linearly to $h = -(-4) - 4 = 0$ m at the groundwater table. The soil profile was discretized into 0.05-m thick elements, which resulted in 81 nodes for the one-dimensional simulation.

18.2.8.1 Single-Layer Soil

The numerically calculated infiltration process for a homogeneous sand is shown in Fig. 18.9 in terms of pressure head ($h - z$) and water content ($\theta - z$) distributions versus depth. At $t < 0$, the initial pressure head is equal to the height above the water table, i.e. $h = 0$ at the water table and $h = -4$ at the soil surface (also see Fig. 18.8). The corresponding water content profile has the same shape as the soil water retention curve. At $t > 0$, water starts to infiltrate and the infiltration front moves downward at a nearly constant rate. After 70 h following the start of the infiltration, the water content in the profile reaches a steady-state value, while a constant water flux is established in the entire soil profile.

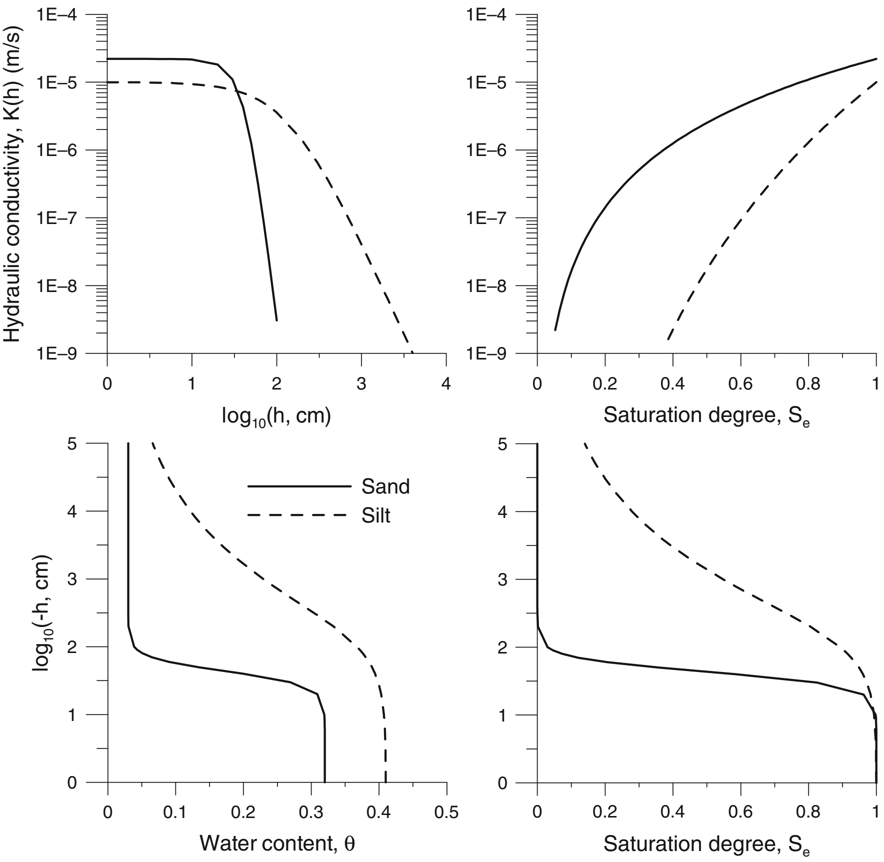


Fig. 18.7 Hydraulic functions for a sand and a silt soil

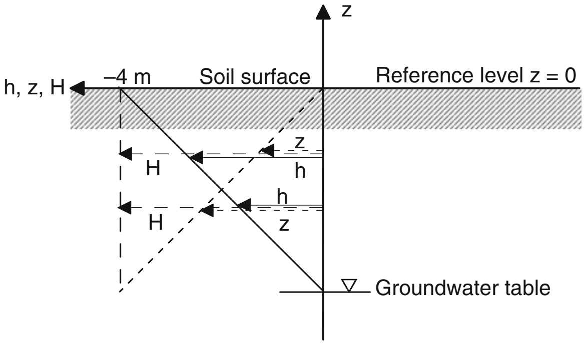


Fig. 18.8 Variation of the components *h* and *z* of the total head *H* along a vertical profile of an unsaturated soil. The total head $H = -4$ m everywhere

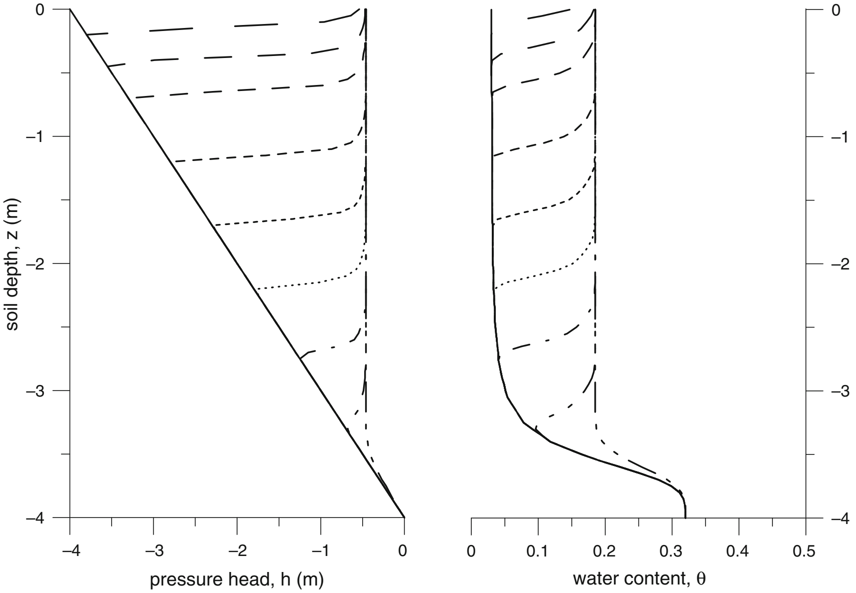


Fig. 18.9 Numerically calculated pressure heads (left) and water contents (right) versus depth for infiltration in a homogeneous sand assuming a water flux q of 2×10^{-6} m/s. The solid line is for $t = 0$, while the distributions at other times (1, 5, 10, 20, 30, 40, 50, 60, 70 h) are indicated by dashes of decreasing length

Having a constant water content between 0 and 3 m would greatly facilitate any sampling effort when the water content has to be measured in a deep homogeneous field soil profile. As water infiltrated, a less negative pressure head developed, reaching a constant value of −0.46 m from the soil surface to close to the water table. In other words, with a pressure head gradient $\mathrm{d}h/\mathrm{d}z \approx 0$, the hydraulic gradient $\mathrm{d}(h+z)/\mathrm{d}z$ becomes one, which corresponds to a unit gradient. The water flux is then determined by gravity only, often referred to as gravity flow. In that case, the unsaturated hydraulic conductivity equals the applied water flux and the water content in the profile is equal to the water content of hydraulic conductivity at that flux.

Results for the homogeneous silt soil using the same water flux of 2×10^{-6} m/s are shown in Fig. 18.10. The flux is now approximately five times lower than the saturated hydraulic conductivity. Compared to the sand, the degree of saturation degree will now be much higher. In contrast to the sand, the initial water content profile is now much smoother, consistent with the reduced nonlinearity of the soil water retention curve. The higher initial (at t=0) degree of saturation for the silt means that a smaller volume of pore space has to be filled with water during the infiltration process as compared to the sand. As a result, the infiltration front moves faster downward in the silt than in the sand.

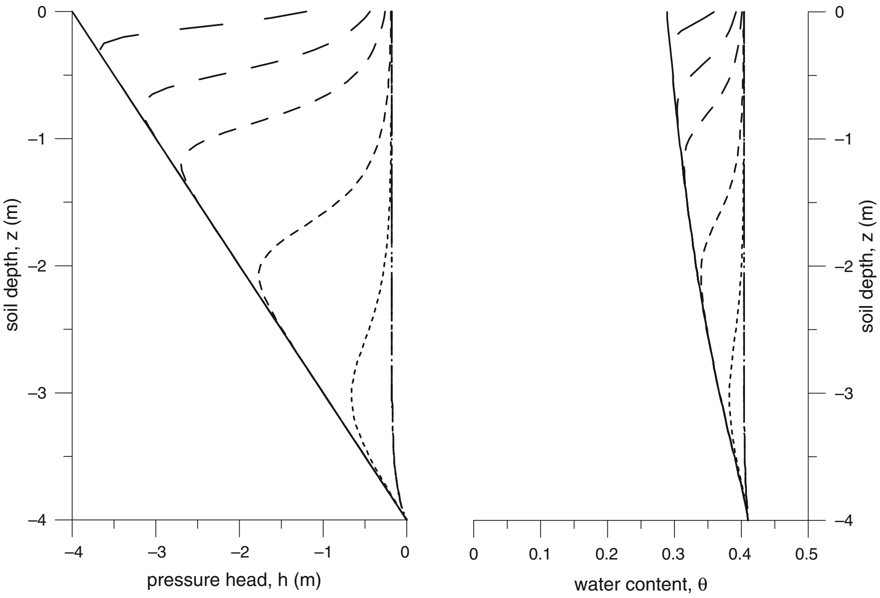


Fig. 18.10 Numerically calculated pressure head (*left*) and water content (*right*) versus depth for a homogeneous silt, assuming a water flux q of 2×10^{-6} m/s. Solid line is for $t = 0$, while the distributions at other times (1, 5, 10, 20, 30, 40, 50, 60, 70 h) are indicated by dashes of decreasing length

18.2.8.2 Two-Layer Soil

Heterogeneous soil profiles are the rule rather than the exception. In general, two or more layers (soil horizons) are present parallel to the surface. Although water flow is not necessarily one-dimensional in such cases, one can often approximate the infiltration process by a one-dimensional model.

The first example considers a sand over silt layer. Hydraulic properties, initial and boundary conditions are those used for the single-layer soil described above. For the water content, the initial condition displays now a discontinuity at the interface between both layers (Fig. 18.11). The pressure head, on the other hand, is continuous across the interface. This behaviour follows from the Buckingham-Darcy equation (Eq. 18.9) which has a finite q only when dh/dz is finite, or when h is continuous everywhere. Up to $t = 30$ h, the infiltration into the sand behaves identically as in the homogeneous sand. The pressure head behind the infiltration front is −0.46 m and the corresponding conductivity in the sand is 2×10^{-6} m/s. Owing to the continuity of the pressure head across the interface, the silt soil has a conductivity of 1×10^{-6} m/s, which is two times smaller than in the sand. As a result, water flow is retarded in the silt and the water content above the interface increases. This in turn leads to a higher K in the silt layer. At steady-state, a unit gradient condition is present

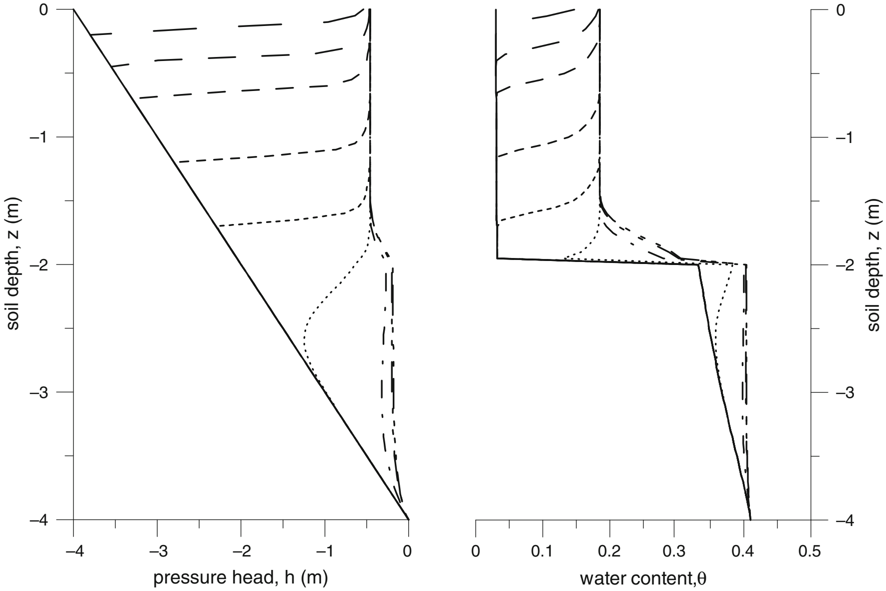


Fig. 18.11 Numerically calculated pressure head (*left*) and water content (*right*) versus depth for infiltration in a two-layer (sand over silt) soil profile, assuming a flux q of 2×10^{-6} m/s. The solid line is for $t = 0$, while distributions at other times (1, 5, 10, 20, 30, 40, 50, 60, 70 h) are indicated by dashes of decreasing length

in both layers, except in a fairly thin zone near the interface. The consequence of having several soil layers on sampling density is obvious from Fig. 18.11: very few samples are required within each homogeneous layer, but the zone above the interface requires a more dense sampling scenario.

18.3 Contaminant Transport

Similarly as Eq. (18.8) for water flow, mathematical formulations for contaminant transport are based on a mass balance equation of the form:

$$\frac{\partial C_T}{\partial t} = -\frac{\partial J_T}{\partial_z} - \phi \tag{18.23}$$

where C_T is the total concentration of contaminant in all forms [ML^{-3}], J_T is the total contaminant mass flux density (mass flux per unit area per unit time) [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$], and ϕ is the rate of change of mass per unit volume by reactions or other sources (negative) or sinks (positive) such as plant uptake [$\mathrm{ML}^{-3}\mathrm{T}^{-1}$].

In its most general interpretation, (Eq. (18.23)) allows the contaminant to exist in all three phases of the soil (i.e., soil gas, liquid phase, and solid phase). The equation further allows for a broad range of transport processes (including advective transport, diffusion, and hydrodynamic dispersion in both liquid phase and the soil gaseous phase), while accommodating any type of chemical reaction that leads to losses or gains in the total concentration.

Whereas the majority of contaminants in soils are usually present only in the liquid phase and solid phase, and as such are transported in the vadose zone mostly only in and by water, some contaminants such as many organic contaminants (e.g., VOCs, BETX), ammonium, fumigants, carbon dioxide, methane, and naturally occurring radioactive gases including radon can have a significant portion of their mass in the gaseous phase, and are thus also subject to transport in the gaseous phase. The total contaminant concentration C_T can be described as (Jury et al. 1991):

$$C_T = \rho_b s + \theta c + a g \tag{18.24}$$

where ρ_b is the bulk density [$\mathrm{ML^{-3}}$], θ is the volumetric water content [$\mathrm{L^3L^{-3}}$], a is the volumetric air content [$\mathrm{L^3L^{-3}}$], and s [$\mathrm{MM^{-1}}$], c [$\mathrm{ML^{-3}}$], and g [$\mathrm{ML^{-3}}$] are concentrations in the solid phase, liquid phase, and soil gas, respectively. The solid phase concentration represents contaminants sorbed onto sorption sites of the solid phase, but can also include contaminants sorbed onto colloids attached to the solid phase or strained by the porous system, and contaminants precipitated onto or into the solid phase.

Various chemical or biological reactions that lead to a loss or gain of contaminant in the soil system may be represented by the reaction term ϕ of Eq. (18.23). This includes radionuclide decay, biological degradation, and/or mineral precipitation-dissolution. In simulation models these reactions are most commonly expressed using zero- and/or first-order reaction rates:

$$\phi = \rho_b s \mu_s + \theta c \mu_w + a g \mu_g - \rho_b \gamma_s - \theta \gamma_w - a \gamma_g \tag{18.25}$$

where μ_s, μ_w, and μ_g are first-order degradation constants for the solid phase, liquid phase, and soil gas [$\mathrm{T^{-1}}$], respectively, and γ_s [$\mathrm{T^{-1}}$], γ_w [$\mathrm{ML^{-3}T^{-1}}$], and γ_g [$\mathrm{ML^{-3}T^{-1}}$] are zero-order production constants for the solid phase, liquid phase, and soil, respectively.

18.3.1 Transport Processes

For contaminants that are present in both liquid phase and soil gas, the total contaminant mass flux, J_T, is the result of various transport processes in both of these phases:

$$J_T = J_l + J_g \tag{18.26}$$

where J_l and J_g represent contaminant fluxes in the liquid phase and soil gas [ML^{-2}T^{-1}], respectively. The three main transport processes that can be active in both liquid phase and soil gas are molecular diffusion, hydrodynamic dispersion, and advection (often also called convection). The contaminant fluxes in the water and gas phases are then the sum of fluxes due to these different processes:

$$\begin{aligned} J_l &= J_{lc} + J_{ld} + J_{lh} \\ J_g &= J_{gc} + J_{gd} + J_{gh} \end{aligned} \tag{18.27}$$

where the subscripts c, d, and h denote convection (or advection), molecular diffusion, and hydrodynamic dispersion, respectively.

18.3.1.1 Diffusion

Diffusion is a result of the random motion of molecules. This process causes a contaminant to move from a location with a higher concentration to a location with a lower concentration. Diffusive transport can be described using Fick's first law:

$$J_d = -\boldsymbol{\theta} D \frac{\partial_c}{\partial_z} = -\boldsymbol{\theta} \boldsymbol{\xi}(\boldsymbol{\theta}) D_o \frac{\partial_c}{\partial_z} \tag{18.28}$$

where D is liquid phase diffusion coefficient [L^{2}T^{-1}], D_o is the diffusion coefficient of the contaminant in free water [L^{2}T^{-1}], and ξ is a tortuosity factor to account for the increased path lengths and decreased cross-sectional areas of the diffusing contaminant in both phases (Jury and Horton 2004).

The parameter D is related to the effective diffusion coefficient D_{eff} (diffusion through the pore space of the porous medium) following $D_{\text{eff}} = D \times \theta$. Values for D_o vary between 2×10^{-9} m^2/s for anions like Cl$^-$ and Br$^-$ and 0.78×10^{-9} m^2/s for cations like Ca^{2+}. Since contaminant diffusion in the liquid phase is severely hampered by both air and solid particles, the tortuosity factor as defined here increases strongly with water content. Many empirical models have been suggested in the literature to account for the tortuosity (e.g., Moldrup et al. 1998). Among these, the most widely used model for the tortuosity factor is probably the equation of Millington and Quirk (1961) given by:

$$\xi(\theta) = \frac{\theta^{7/3}}{\theta_s^2} \tag{18.29}$$

where θ_s is the saturated water content [L^{3}L^{-3}].

Jin and Jury (1996) suggested that the following version of the Millington-Quirk model may be in better agreement with experimental data:

$$D = \frac{\theta}{\theta_s^{2/3}} D_o \tag{18.30}$$

As an example, taking $\theta_s = 0.37$ and $D_o = 10^{-9}$ m^2/s, and assuming saturated or nearly saturated conditions (i.e. $\theta = \theta_s = 0.37$), then $D = 0.72 \times 10^{-9}$ m^2/s.

18.3.1.2 Dispersion

Dispersive transport of contaminants results from the heterogeneous distribution of water flow velocities within and between different soil pores (Fig. 18.12, left). Dispersion can be derived from Newton's law of viscosity which states that velocities within a single capillary tube follow a parabolic distribution, with the largest velocity in the middle of the pore and zero velocities at the walls (Fig. 18.12, left). For this reason contaminants in the middle of a pore will travel faster than contaminants that are farther from the center. Since the distribution of contaminant ions within a pore depends on their charge, as well as on the charge of pore walls, some contaminants may move considerably faster than others. In some circumstances (e.g., in fine-textured soils), anion exclusion may occur. Since the anions in such soils are located predominantly in the faster moving liquid in the center of a pore away from the negatively charge solid surface, anionic contaminants may travel faster than the average velocity of water (e.g., Nielsen et al. 1986). Based on Poiseuille's law, one can further show that velocities in a capillary tube depend strongly on the radius of the tube, and that the average velocity increases with the radius to the second power. Since soils consist of pores of many different diameters, contaminant fluxes in pores of different diameters will be significantly different, with some contaminants again traveling faster than others (Fig. 18.12, right). Furthermore, contaminants may travel according to pathways of different length. All these factors result in a bell-shaped distribution of velocities and thus of arrival times, typical of a breakthrough curve.

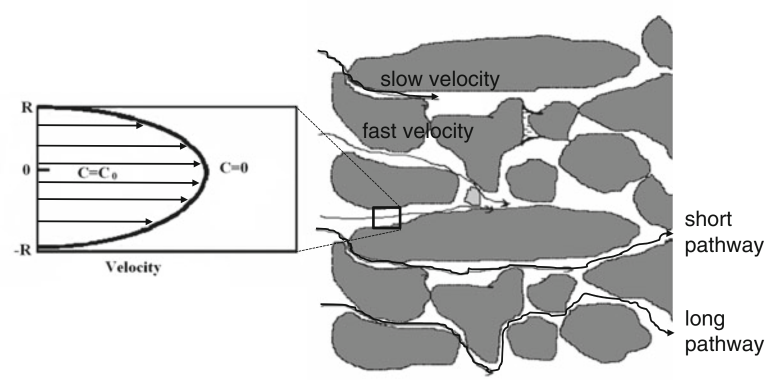


Fig. 18.12 Distribution of velocities in a single pore (*left*) and distribution of velocities in a pore system (*right*) resulting in different arrival times of contaminants (modified from Šimůnek and Van Genuchten 2006)

The above pore-scale dispersion processes lead to an overall (macroscopic) hydrodynamic dispersion process that can be described mathematically in approximately the same way as molecular diffusion using Fick's first law. Adding the dispersion and diffusion processes leads then to:

$$J_h = -\boldsymbol{\theta} D_h \frac{\partial c}{\partial c} = -\boldsymbol{\theta}(D_m + D)\frac{\partial c}{\partial z} \tag{18.31}$$

where D_h is the hydrodynamic dispersion coefficient [L^2T^{-1}] that accounts for both molecular diffusion and mechanical dispersion (Fetter 1999), D_m is the mechanical dispersion coefficient [L^2T^{-1}], and D is the liquid phase diffusion coefficient [L^2T^{-1}]. The mechanical dispersion coefficient in one-dimensional systems has been found to be approximately proportional to the average pore-water velocity v ($= q/\theta$) [LT^{-1}], with the proportionality constant generally referred to as the (longitudinal) dispersivity λ (Biggar and Nielsen 1967). The discussion above holds for one-dimensional transport; multi-dimensional applications require the use of a more complicated dispersion tensor involving longitudinal and transverse dispersivities (e.g., Bear 1972).

Dispersivity is a transport parameter that is often obtained experimentally by fitting measured breakthrough curves with analytical solutions of the advection-dispersion equation (discussed further below). The dispersivity often changes with the distance over which contaminants travel. Values of the longitudinal dispersivity usually range from about 1 cm for relatively short, packed laboratory columns, to about 5 or 10 cm for field soils. Longitudinal dispersivities can be significantly larger (up to hundreds of meters) for regional groundwater transport problems (Gelhar et al. 1985). If no other information is available, a good first approximation is to use a value of one-tenth of the transport distance for the longitudinal dispersivity (e.g, Anderson 1984), and a value of one-hundred of the transport distance for the transverse dispersivity when multi-dimensional applications are considered.

18.3.1.3 Advection

Advective transport refers to contaminants being transported with the moving fluid, either in the liquid phase (J_{lc}) or the soil gas (J_{gc}), i.e.:

$$J_c = qc \tag{18.32}$$

where we now use a more general notation by omitting the subscripts l and g. Advective transport generally is dominant in the liquid phase. Advective transport may also occur in the gaseous phase, but is often neglected since its contribution is generally negligible compared to gaseous diffusion.

The total contaminant flux density in both the liquid phase and soil gas is obtained by incorporating contributions from the various transport processes into Eq. (18.27) to obtain

$$J = qc - \theta D_h \frac{\partial c}{\partial z} \tag{18.33}$$

where D_h is the hydrodynamic dispersion coefficient [L^2T^{-1}] that accounts for both molecular diffusion and mechanical dispersion.

Mechanical dispersion in most subsurface transport problems dominates molecular diffusion in the liquid phase, except when the fluid velocity becomes relatively small or is negligible. Diffusion dominated transport occurs in low permeability media, such as clays, rock matrices, and man-made structures such as concrete.

18.3.2 Advection-Dispersion Equations

18.3.2.1 Transport Equations

The equation governing transport of dissolved contaminants in the vadose zone is obtained by combining the contaminant mass balance (Eq. (18.23)) with equations defining the total concentration of the contaminant (Eq. (18.24)) and the contaminant flux density (Eq. (18.33)) to give

$$\frac{\partial(\rho_b s + \theta c + ag)}{\partial t} = \frac{\partial}{\partial z}(\theta D_h \frac{\partial c}{\partial z}) + \frac{\partial}{\partial z}(aD_g^s \frac{\partial g}{\partial z}) - \frac{\partial(qc)}{\partial z} - \phi \tag{18.34}$$

where D_h and ${D^s}_g$ are the hydrodynamic dispersion coefficient in the liquid and gaseous phases [L^2T^{-1}], respectively.

Several alternative formulations of Eq. (18.34) can be found in the literature. For example, for one-dimensional transport of non-volatile contaminants, Eq. (18.34) simplifies to

$$\frac{\partial(\rho_b s + \theta c)}{\partial t} = \frac{\partial(\theta R c)}{\partial t} = \frac{\partial}{\partial z}\left(\theta D_h \frac{\partial c}{\partial z}\right) - \frac{\partial(qc)}{\partial z} - \phi \tag{18.35}$$

where q is the vertical water flux density [LT^{-1}] and R is the retardation factor [−]:

$$R = 1 + \frac{\rho_b}{\theta}\frac{ds(c)}{dc} \tag{18.36}$$

For transport of inert, non-adsorbing contaminants during steady-state water flow a further simplification is possible:

$$\frac{\partial c}{\partial t} = D_h \frac{\partial^2 c}{\partial z^2} - v\frac{\partial c}{\partial z} \tag{18.37}$$

The above equations are usually referred to as advection-dispersion equations (ADEs), or convection-dispersion equations (CDEs).

18.3.2.2 Linear and Non-Linear Sorption

The advection-dispersion equation given by Eq. (18.34) contains three unknown concentrations (those for the soil phase, liquid phase, and soil gas), while Eq. (18.35) contains two unknowns. To be able to solve these equations, additional information is needed that relates solid phase and liquid phase concentrations to each other. A common approach is to assume instantaneous sorption and to use adsorption isotherms to relate the liquid and adsorbed concentrations. The simplest form of the adsorption isotherm is a linear isotherm, sometimes very appropriate at low contaminant concentrations, given by:

$$s = K_d c \tag{18.38}$$

where K_d is the distribution or partitioning coefficient [L^3M^{-1}]. Equation (18.38) assumes reversible sorption (adsorption equals desorption). Substitution of this equation into Eq. (18.36) leads to a constant value for the retardation factor, i.e.,

$$R = 1 + \frac{\rho K_d}{\theta} \tag{18.39}$$

Whereas the use of a linear isotherm greatly simplifies the mathematical description of contaminant transport, sorption and exchange are generally non-linear and most often depend also on the presence of competing species in the liquid phase. Unlike linear adsorption, the contaminant retardation factor for non-linear adsorption is not constant, but changes as a function of concentration. Many models have been used in the past to describe non-linear sorption. Two widely used non-linear sorption models are those by Freundlich (1909) and Langmuir (1918) given by:

$$s = K_f c^\beta \tag{18.40}$$

$$s = \frac{K_d c}{1 + \eta c} \tag{18.41}$$

respectively, where K_f [$L^{3\beta}$ $M^{-\beta}$] and β [$-$] are coefficients in the Freundlich isotherm, and η [L^3M^{-1}] is a coefficient in the Langmuir isotherm. Examples of linear, Freundlich and Langmuir adsorption isotherms are given in Fig. 18.13.

18.3.2.3 Volatilization

Volatilization is increasingly recognized as an important process affecting the fate of many organic contaminants, including pesticides, fumigants, and explosives in soils (Jury et al. 1983a, 1984; Glotfelty and Schomburg 1989). While many organic contaminants dissipate by means of chemical and microbiological degradation, volatilization may be equally important for volatile substances, such as certain pesticides. The volatility of pesticides is influenced by many factors, including the

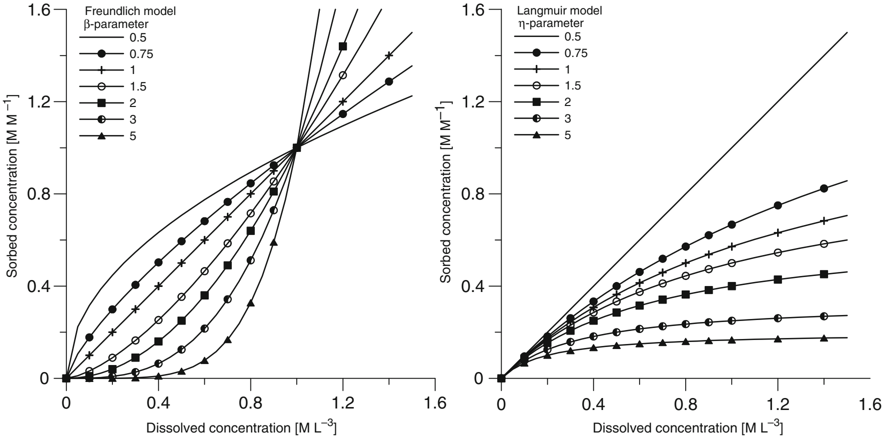


Fig. 18.13 Plots of the Freundlich adsorption isotherm given by Eq. (18.40), with $K_f = 1$ and β defined in the caption (*left*), and the Langmuir adsorption isotherm given by Eq. (18.41), with $K_d = 1$ and η defined in the caption (*right*)

physicochemical properties of the contaminant itself as well as such environmental variables as temperature and solar energy. Even though only a small fraction of a pesticide may exist in the gas phase, air-phase diffusion rates can sometimes be comparable to diffusion in the liquid phase since gaseous diffusion coefficients are about 4 orders of magnitude greater than liquid phase diffusion coefficients.

The general transport equation given by Eq. 18.34 can be simplified considerably when assuming linear equilibrium sorption and volatilization such that the adsorbed (s) and gaseous (g) concentrations are linearly related to the solution concentration (c) through the distribution coefficients K_d (Eq. (18.38)) and K_H, the latter appearing in

$$g = K_H c \tag{18.42}$$

where K_H is the dimensionless Henry constant [−].

Assuming linear partitioning, Eq. (18.34) for one-dimensional transport then has the form

$$\frac{\partial(\rho_b K_d + \theta + aK_H)c}{\partial t} = \frac{\partial}{\partial z}\left(\theta D_h \frac{\partial c}{\partial z}\right) + \frac{\partial}{\partial z}\left(aD_g^s K_H \frac{\partial c}{\partial z}\right) - \frac{\partial(qc)}{\partial x} - \phi \tag{18.43}$$

or

$$\frac{\partial \theta R c}{\partial t} = \frac{\partial}{\partial z}\left(\theta D_E \frac{\partial c}{\partial z}\right) - \frac{\partial(qc)}{\partial x} - \phi \tag{18.44}$$

where the retardation factor R [−] and the effective dispersion coefficient D_E [L^{2}T^{-1}] are defined as follows

$$R = 1 + \frac{\rho_b K_d + a K_H}{\theta}$$
$$D_E = D_h + \frac{a D_g^s K_H}{\theta} \tag{18.45}$$

Jury et al. (1983b, 1984) provided for many organic contaminants their distribution coefficients K_d and Henry's constants K_H, from which calculated percent mass present in each phase may be obtained.

18.3.2.4 Initial and Boundary Conditions

Initial Conditions

The governing equations for one-dimensional contaminant transport can be solved analytically or numerically if the initial and boundary conditions for the problem being considered are specified. Initial conditions need to be provided for one equilibrium phase concentration, i.e.,

$$c(z,t) = c_i(z,0) \tag{18.46}$$

where c_i is the initial concentration [ML^{-3}], as well as for all non-equilibrium phases if nonequilibrium transport is to be considered (see further).

Boundary Conditions

Complex interactions between the transport domain and its environment often must be considered for the water flow part of the problem at hand since these interactions determine the magnitude of water fluxes across the domain boundaries. By comparison, the contaminant transport part of most analytical and numerical models usually considers only three types of boundary conditions. When the concentration at the boundary is known, one can use a first-type (or Dirichlet type) boundary condition of the form:

$$c(z,t) = c_0(z,t) \tag{18.47}$$

where c_0 is a prescribed concentration [ML^{-3}] at the top or bottom of the soil profile. This boundary condition is often referred to as a concentration boundary condition. A third-type (Cauchy type) boundary condition may be used to prescribe the concentration flux at the boundary as follows:

$$-\boldsymbol{\theta} D\frac{\partial c}{\partial x} + qc = qc_0 \tag{18.48}$$

in which q represents the boundary fluid flux [LT^{-1}], and c_0 is the concentration of the incoming fluid [ML^{-3}]. In some cases, for example when a boundary is impermeable (q=0) or when water flow is directed out of the region, the Cauchy boundary condition reduces to a second-type (Neumann type) boundary condition of the form:

$$\theta D \frac{\partial c}{\partial x} = 0 \tag{18.49}$$

Most applications require a Cauchy boundary condition rather than Dirichlet (or concentration) boundary condition. Since Cauchy boundary conditions define the contaminant flux across a boundary, the contaminant flux entering the transport domain will be known exactly (as specified). This specified contaminant flux is then in the transport domain divided into advective and dispersive components. On the other hand, a Dirichlet boundary condition controls only the concentration on the boundary, but not the contaminant flux into the domain, which because of its advective and dispersive contributions will be larger than for a Cauchy boundary condition. The incorrect use of Dirichlet rather than Cauchy boundary conditions may lead to significant mass balance errors at early times, especially for relative short transport domains (Van Genuchten and Parker 1984).

18.3.3 Nonequilibrium Transport

Because equilibrium contaminant transport models have frequently been unable to describe experimental data accurately, a large number of diffusion-controlled physical nonequilibrium and chemical-kinetic models have been developed and used to describe the transport of both non-adsorbing and adsorbing contaminants (Šimůnek and Van Genuchten 2008). Efforts to model nonequilibrium transport usually involve relatively simple first-order rate equations. Nonequilibrium models have used the assumptions of two-region (dual-porosity) type transport involving contaminant exchange between mobile and immobile liquid phase transport regions, and/or one-, two- or multi-site sorption formulations (e.g., Brusseau 1999; Nielsen et al. 1986). Models simulating the transport of particle-type contaminants, such as colloids, viruses, and bacteria, often also use first-order rate equations to describe such processes as attachment, detachment, and straining. In many cases nonequilibrium models have resulted in better descriptions of measured laboratory and field contaminant transport data, in part by providing additional degrees of freedom for fitting measured concentration distributions.

18.3.3.1 Physical Nonequilibrium

Dual-Porosity and Mobile-Immobile Water Models

Two-region transport models (Fig. 18.14b and c) assume that the liquid phase can be partitioned into distinct mobile (flowing) and immobile (stagnant) liquid pore regions, and that contaminant exchange between the two liquid regions can be modeled as a first-order exchange process. Using the same notation as before, the two-region contaminant transport model is given by (Toride et al. 1993; Van Genuchten and Wagenet 1989):

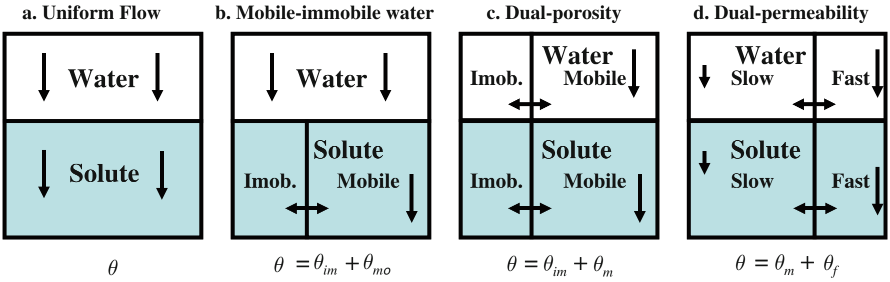


Fig. 18.14 Conceptual models of water flow and contaminant transport (θ is the water content, θ_{mo} and θ_{im} in (**b**) and (**c**) are water contents of the mobile and immobile flow regions, respectively, with $\theta_{mo}/(\theta_{mo} + \theta_{im}) = \phi_m$, and θ_m and θ_f in (**d**) are water contents of the matrix and macropore (fracture) regions, respectively) (after Šimůnek and Van Genuchten (2006))

$$\begin{aligned}\frac{\partial \theta_{mo} c_{mo}}{\partial t} + \frac{\partial f \rho s_{mo}}{\partial t} &= \frac{\partial}{\partial z}\left(\theta_{mo} D_{mo} \frac{\partial c_{mo}}{\partial z}\right) - \frac{\partial q c_{mo}}{\partial z} - \phi_{mo} - \Gamma_s \\ \frac{\partial \theta_{im} c_{im}}{\partial t} + \frac{\partial (1-f) \rho s_{im}}{\partial t} &= -\phi_{im} + \Gamma_s\end{aligned} \tag{18.50}$$

for the mobile (macropores, subscript *mo*) and immobile (matrix, subscript *im*) domains, respectively, where f is the dimensionless fraction of sorption sites in contact with the mobile water [−], ϕ_{mo} and ϕ_{im} are reactions in the mobile and immobile domains [ML^{3}T^{-1}], respectively, and Γ_s is the contaminant transfer rate between the two regions [ML^{3}T^{-1}]. The same Eq. (18.50) can be used to describe contaminant transport considering both the mobile-immobile and dual-porosity models shown in Fig. 18.14b and c, respectively.

Dual-Permeability Model

One approach for implementing a dual-permeability formulation for contaminant transport is to assume advection-dispersion type equations for transport in both the fracture and matrix regions as follows (Gerke and Van Genuchten 1993):

$$\frac{\partial \theta_f c_f}{\partial t} + \frac{\partial \rho s_f}{\partial t} = \frac{\partial}{\partial z}\left(\theta_f D_f \frac{\partial c_f}{\partial z}\right) - \frac{\partial q_f c_f}{\partial z} - \phi_f - \frac{\Gamma_s}{w} \tag{18.51}$$

$$\frac{\partial \theta_m c_m}{\partial t} + \frac{\partial \rho s_m}{\partial t} = \frac{\partial}{\partial z}\left(\theta_m D_m \frac{\partial c_m}{\partial z}\right) - \frac{\partial q_m c_m}{\partial z} - \phi_m - \frac{\Gamma_s}{1-w} \tag{18.52}$$

where the subscript f and m refer to the macroporous (fracture) and matrix pore systems, respectively; ϕ_f and ϕ_m represent sources or sinks in the macroporous and matrix domains [ML^{3}T^{-1}], respectively; and w is the ratio of the volume of the macropore (inter-aggregate) domain and that of the total soil system [-]. Equations (18.51) and (18.52) assume advective-dispersive type transport descriptions for both the fractures and the matrix. Several authors simplified transport in the macropore domain, for example by ignoring diffusion and dispersion in the macropores and considering only piston displacement there (Ahuja and Hebson 1992; Jarvis et al. 1994).

Mass Transfer

The transfer rate, Γ_s, in Eq. (18.50) for contaminants between the mobile and immobile domains in the dual-porosity models can be given as the sum of diffusive and advective fluxes as follows:

$$\Gamma_s = \alpha_s(c_{mo} - c_{im}) + \Gamma_w c^* \tag{18.53}$$

where c^* is equal to c_{mo} for $\Gamma_w > 0$ and c_{im} for $\Gamma_w < 0$, and α_s is the first-order contaminant mass transfer coefficient [T^{-1}]. Notice that the advection term of Eq. 18**.**53 is equal to zero for the mobile-immobile model (Fig. 18.14b) since the immobile water content in this model is assumed to be constant. However, Γ_w may have a nonzero value in the dual-porosity model depicted in Fig. 18.14c.

The transfer rate, Γ_s, in Eqs. (18.51) and (18.52) for contaminants between the fracture and matrix regions is also usually given as the sum of diffusive and advective fluxes as follows (e.g., Gerke and Van Genuchten 1996):

$$\Gamma_s = \alpha_s(1 - w_m)(c_f - c_m) + \Gamma_w c^* \tag{18.54}$$

in which the mass transfer coefficient, Γ_s [T^{-1}], is of the form:

$$\alpha_s = \frac{\beta_g}{d^2} D_a \tag{18.55}$$

where β_g is a dimensionless geometry-dependent coefficient, d is the characteristic length of the matrix structure (L) (e.g., the radius of a spherical or solid cylindrical aggregate, or half the fracture spacing in the case of parallel rectangular voids), D_a is an effective diffusion coefficient [L^{2}T^{-1}] representing the diffusion properties of the fracture-matrix interface.

18.3.3.2 Chemical Nonequilibrium

Kinetic Sorption Models

A substitute to expressing sorption as an instantaneous process using algebraic equations (e.g., Eqs. (18.38), (18.40) or (18.41)) is to quantify the reaction kinetics based on ordinary differential equations. A popular and simple formulation of a chemically controlled kinetic reaction arises when first-order linear kinetics is assumed:

$$\frac{\partial s}{\partial t} = \alpha_k(K_d c - s) \tag{18.56}$$

where α_κ is a first-order kinetic rate coefficient [T^{-1}]. Several other nonequilibrium adsorption expressions have been used in the past (see Table 2 in Van Genuchten and Šimůnek 1996), and are often referred to as one-site sorption models.

Because contaminant transport models assuming chemically controlled nonequilibrium (one-site sorption) often did not lead to significant improvements in predictive capability when used to describe column transport experiments, the one-site first-order kinetic model was extended into a two-site sorption concept that divides the available sorption sites into two fractions (Selim et al. 1976; Van Genuchten and Wagenet 1989). Conceptually, sorption on one fraction (type-1 sites) is assumed to be instantaneous, while sorption on the remaining (type-2) sites is considered to be time-dependent. Assuming linear sorption, the two-site transport model is given by (Van Genuchten and Wagenet 1989):

$$\begin{aligned} \frac{\partial(f\rho_b K_d + \theta)c}{\partial t} &= \frac{\partial}{\partial z}\left(\theta D_h \frac{\partial c}{\partial z}\right) - \frac{\partial(qc)}{\partial z} - \phi_e \\ \frac{\partial s_k}{\partial t} &= \alpha_k[(1-f)K_d c - s_k] - \phi_k \end{aligned} \tag{18.57}$$

where f is the fraction of exchange sites assumed to be at equilibrium [−], ϕ_e [$\mathrm{ML}^3\mathrm{T}^{-1}$] and ϕ_k [$\mathrm{MM}^{-1}\mathrm{T}^{-1}$] are reactions in the equilibrium and nonequilibrium phases, respectively, and the subscript k refers to kinetic (type-2) sorption sites. When $f = 0$, the two-site sorption model reduces to the one-site fully kinetic sorption model (i.e., when only type-2 kinetic sites are present). However, if $f = 1$, the two-site sorption model reduces to the equilibrium sorption model for which only type-1 equilibrium sites are present.

18.3.3.3 Colloid-Facilitated Solute Transport

There is considerable evidence that many contaminants, including radionuclides (Noell et al. 1998; Von Gunten et al. 1988), pesticides (Kan and Tomson 1990; Lindqvist and Enfield 1992; Vinten et al. 1983), heavy metals (Grolimund et al. 1996), viruses, pharmaceuticals (Thiele-Bruhn 2003; Tolls 2001), hormones (Hanselman et al. 2003), and other contaminants (Magee et al. 1991; Mansfeldt et al. 2004) in the subsurface are transported not only with moving water, but also sorbed to mobile colloids. Because many colloids and microbes are negatively charged and thus electrostatically repelled by negatively-charged solid surfaces, the process of anion exclusion may occur. As a result, contaminant transport could be slightly enhanced relative to water flow. Size exclusion may similarly enhance the advective transport of colloids by limiting their presence and mobility to the larger pores (e.g., Bradford et al. 2003). Sorption of contaminants onto mobile colloids can thus significantly accelerate their transport relative to more standard advection-transport circumstances.

Colloid-facilitated transport is a relatively complicated process that requires knowledge of water flow, colloid transport, dissolved contaminant transport, and colloid-contaminant interaction. This requires formulation of transport and/or mass balance equations for both water flow and colloid transport, and in addition for the total contaminant, for contaminants sorbed kinetically or instantaneously to the solid phase, and for contaminants sorbed to mobile colloids, to colloids attached to the soil solid phase, and to colloids accumulating at the air-water interface. Development of such a model is beyond the scope of this chapter. The interested reader will find additional documentation from de Jonge et al. (2004), Flury and Qiu (2008), Massoudieh and Ginn (2009), Šimůnek et al. (2006a), and Totsche and Kögel-Knabner (2004).

18.3.4 Stochastic Models

Several studies have demonstrated that solutions of classical contaminant transport models often fail to accurately describe transport processes in most field soils. A major reason for the inability to capture field-scale contaminant transport is the fact that the subsurface environment is very heterogeneous. Heterogeneity occurs at different levels of spatial and time scales (Wheatcraft and Cushman 1991), ranging from microscopic scales involving time-dependent chemical sorption and precipitation/dissolution reactions, to intermediate scales involving the preferential movement of water and contaminants through macropores or fractures, and further to large scales involving the spatial variability of soils across the landscape. Soil heterogeneity can be addressed in terms of process-based descriptions which attempt to consider the effects of heterogeneity at one or several scales. It can also be tackled using stochastic approaches which incorporate certain assumptions about the transport process in the heterogeneous system (e.g., Dagan 1989; Sposito and Barry 1987). In this Section we briefly review flow and transport parameter heterogeneity and discuss several stochastic transport approaches, notably those using stream tube models and the transfer function approach.

18.3.4.1 Flow and Transport Parameter Heterogeneity

Spatial Variation of Physical and Chemical Properties

Hydrological and geological processes are known to vary in space (Delhomme 1979; Nielsen et al. 1973). This is due to the natural evolution of the subsurface over geological time scales and anthropogenic activities in recent times. Knowledge of spatial variability of physical and chemical properties is of utmost importance not only from the point of view of characterisation of the heterogeneous subsurface, but also in the context of long term prediction of water flow and contaminant transport in the subsurface. In this context, the last four decades has seen significant number of studies carried out in order to comprehend the spatial variability of physical and chemical properties of soil with the help of both conventional statistics and geostatistics.

From the vast body of literature, a wealth of data have been meticulously compiled by Jury (1985) and Thibault et al. (1990) on the spatial variability of physical and chemical properties respectively; salient data from these compilations are presented in Tables 18.2 and 18.3. Note that the properties which are treated statistically

Table 18.2 Replicate measurements of field transport parameters and their variability expressed by the coefficient of variation CV

Parameter	CV (%)	Soil texture	Number of observations	References
Static properties				
Porosity	9.6	Loamy sand	12	Cassel (1983)
Porosity	11.4	Clay loam	120	Russo and Bresler (1981)
Bulk density	2.3	Sand	5	Babalola (1978)
Bulk density	26.0	Sandy loam	30	Courtin et al. (1983)
Water content (0.1 bar, low suction)	4.3	Clay loam	100	Gajem et al. (1981)
Water content (0.1 bar, low suction)	20.0	Sandy loam	64	Gumaa (1978)
Water content (15 bar, high suction)	14	Clay loam	900	Gajem et al. (1981)
Water content (15 bar, high suction)	45	Sandy loam	172	Cassel and Bauer (1975)
Water content	7.0	Sandy loam	180	Mallants et al. (1996c)
Water content	23.0	Sandy loam	180	Mallants et al. (1996c)
Van Genuchten parameter, θ_r	54.9–156.4	Sandy loam	180	Mallants et al. (1996c)
Van Genuchten parameter, θ_s	5.1–7.6	Sandy loam	180	Mallants et al. (1996c)
Van Genuchten parameter, α	45.0–53.0	Sandy loam	180	Mallants et al. (1996c)
Van Genuchten parameter, n	11.0–22.0	Sandy loam	180	Mallants et al. (1996c)
Soil pH	2.0	Clay loam	100	Gajem et al. (1981)
Soil pH	15.0	Clay loam	1040	Cameron et al. (1971)
K_d	31.0	Loamy sand	36	El Abd (1984)
CEC	35.5	Loamy clay	93	Jacques et al. (1999)
Langmuir parameter, K_f	101.19	Clay loam	93	Jacques et al. (1999)
Langmuir parameter, n	8.06	Clay loam	93	Jacques et al. (1999)
Sand, Loam, Clay and OC	82.6, 37.2, 32.4 and 78.1	Clay loam	93	Jacques et al. (1999)
Sand	3.0	Loamy sand	5	Babalola (1978)
Clay	55.0	Clay	480	Nielsen et al. (1973)

Table 18.2 (continued)

Parameter	CV (%)	Soil texture	Number of observations	References
Dynamic properties				
K_s	48	Silty clay loam	33	Willardson and Hurst (1965)
K_s	320	Very coarse	36	Willardson and Hurst (1965)
K_s	322 (C1 horizon) and 897 (C2 horizon)	Sandy loam	180	Mallants et al. (1996c)
Water diffusivity	284 (90% saturation), 412 (60%)	Clay loam	20	Nielsen et al. (1973)
Solute concentration	61	Loamy sand	32	Wild and Babiker (1976)
Solute concentration	127	Loamy sand	36	Jury et al. (1983b)
Solute travel time	36	Sandy	14 sites	Jury et al. (1982)
Solute travel time	194	Loamy clay	20 plots	Biggar and Nielsen (1976)

Partly extracted from Jury (1985); extreme values have been presented

Table 18.3 K_d variability based on statistical analysis of literature data (excerpt from Thibault et al. 1990)

Parameter	σ*	Soil texture	Range of K_d values – L Kg^{-1}	Number of observations
Cd	1.6	Silt	7–962	8
	1.5	Sand	2.7−625	14
	0.9	Clay	112−2,450	10
Cr	2.9	Silt	2.2−1,000	4
	2.1	Sand	1.7−1,729	15
	–	Clay	1,500 (exp(μ))	–
Pb	1.4	Silt	3,500–59,000	3
	2.3	Sand	19–1,405	3
	–	Clay	550 (exp(μ))	–
Zn	2.4	Silt	3.6–11,000	12
	2.6	Sand	0.1–8,000	22
	1.4	Clay	200–100,000	23

*Standard deviation of the logarithms of the observed values

in this section can be loosely categorized as *static properties* and dynamic *properties*. Static properties include porosity, bulk density, compositional properties such as particle size fractions, water retention properties, soil pH, cation exchange capacity (CEC) and such distribution coefficients as the linear K_d, the Freundlich K_f and the Langmuir K_d. Dynamic properties include the soil water diffusivity, hydraulic conductivity, contaminant breakthrough curves, dispersivity and contaminant concentrations (Jury 1985). The following sections summarize salient aspects of some of the static and dynamic properties from these compilations, supplemented with recent data on spatial variability.

Static Properties

Jury's (1985) compilations suggest that the range of the coefficient of variation (CV=standard deviation/mean) for static properties is surprisingly small given the variety of soils and field sizes where the experiments were conducted. For instance, the CV for porosity ranged between 9.6–11.4% and soil bulk density varied between 2.3–26%. In some studies, replicate measurements of the volumetric or gravimetric water content in equilibrium with an imposed suction of 0.1 bar showed CV's ranging from 4.3 to 20%; however, when the imposed suction was raised to 15 bars, the CV was higher and had a range of 14–45%. This is similar to observations reported by Mallants et al. (1996c) who obtained a CV of 23% at 2.3 pF compared to 7% at 0 pF, although the trend reversed at higher pFs.

From Jury's (1985) compilations, soil pH was generally found to be quite uniform with a maximum CV of 15%. For the K_d values they reported only one study based on batch and column measurements, which showed a low CV of 31% on soil samples taken across a sandy field of 0.64 ha.

Thibault et al. (1990) presented a comprehensive compilation of K_d values for various soil textures and elements. Table 18.3 only presents the values for some heavy metals for three soil textures (the total number of elements reported by Thibault et al. (1990) was 49). It is evident that the range of values encountered for various soil types varies significantly and over several orders of magnitude. In general, K_d values for sand are lower than for silt or clay for the elements shown. Although not shown in Table 18.3, Thibault et al. (1990) noted that elements such as silver, nickel, americium and thallium showed higher mean K_d values for organic soils than for clay.

In yet another study, Jacques et al. (1999) reported the spatial variability of the Freundlich K_f and n parameters for atrazine and their correlation with soil textural variables, cation exchange capacity and organic carbon content of a stagnic podzoluvisol based on 93 samples. The CV values (of non-transformed variables) for sand, loam, clay and organic carbon content were 82.6, 37.2, 32.4, and 78.1%, respectively. The CV values for CEC, K_f and n were reported to be 35.5, 101.2, and 8.06% respectively. These properties therefore have CV values that vary from moderate to high, with K_f showing the highest variation.

Dynamic Properties

Jury (1985) noted that, in general, the CV values of dynamic properties are higher than those of the static properties. For instance, based on thirteen studies reporting replicated measurements of the saturated K_s, the CV was found to be between 48 and 320%. Interestingly, both the lowest (48%) and the second highest (190%) CV's were measured on finer-textured soils. Thus, Jury (1982) commented that there appears to be no apparent relationship between the variability of K_s and soil type, at least within the soil textural classes considered in their compilation. His findings have been contradicted by others (e.g. Cosby et al. 1984).

As regards the contaminant transport properties, observations of contaminant concentration variability have also been compiled by Jury (1985). The results based on four field scale studies in which enough replicates were taken to allow a meaningful determination of sample concentration variance revealed a range of 61–127% (Table 18.2). He also reported the log variance of contaminant velocity as another transport property of interest. Based on six studies, a relatively low CV of 36% and a high CV of 194% were reported.

18.3.4.2 Stream Tube Models

The downward movement of contaminants from the soil surface to an underlying aquifer may be described stochastically by viewing the field as a series of independent vertical columns, often referred to as "stream tubes" (Fig. 18.15), while contaminant mixing between the stream tubes is assumed to be negligible. Transport in each tube may be described deterministically with the standard advection-dispersion equation (ADE), or modifications thereof to include additional geochemical and microbiological processes. Transport at the field scale is then implemented by considering the column parameters as realizations of a stochastic process, having a random distribution (Toride et al. 1995).

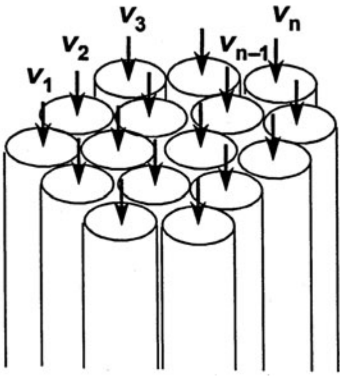


Fig. 18.15 Schematic illustration of the stream tube model (*left*), (Toride et al. 1995)

The stream tube model was implemented into the CXTFIT 2.0 code (Toride et al. 1995) for a variety of transport scenarios in which the pore water velocity in combination with either the hydrodynamic dispersion coefficient, D_h, the distribution coefficient for linear adsorption, K_d, or the first-order rate coefficient for nonequilibrium adsorption, α_k, are stochastic variables (Toride et al. 1995).

18.3.5 Multicomponent Reactive Solute Transport

The various mathematical descriptions of contaminant transport presented thus far all considered contaminants that would move independently of other contaminants in the subsurface. In reality, the transport of reactive contaminants is more often than not affected by interactive physico-chemical and even biochemical processes. Simulating these processes requires a more comprehensive approach that couples the physical processes of water flow and advective-dispersive transport with a range of biogeochemical processes. The liquid phase is always a mixture of many ions which may be involved in mutually dependent chemical processes, such as complexation reactions (Lichtner 1996; Yeh and Tripathi 1990), cation exchange (White and Zelazny 1986), precipitation-dissolution (Šimůnek and Valocchi 2002), sorption-desorption, volatilization, redox reactions, and degradation, among other reactions (Šimůnek and Valocchi 2002). Transport and transformation of many contaminants is further mediated by subsurface aerobic or anaerobic bacteria. Bacteria catalyze redox reactions in which organic compounds (e.g., hydrocarbons) act as the electron donor and inorganic substances (oxygen, nitrate, sulfate, or metal oxides) as the electron acceptor. By catalyzing such reactions, bacteria gain energy and organic carbon to produce new biomass. These and related processes can be simulated using integrated reactive transport codes that couple the physical processes of water flow and advective-dispersive contaminant transport with a range of biogeochemical processes (Jacques et al. 2003; Šimůnek et al. 2006b).

Once the various chemical reactions are defined, the final system of governing equations usually consists of several partial differential equations for contaminant transport (i.e., advection-dispersion equations for each component) plus a set of nonlinear algebraic and ordinary differential equations describing the equilibrium and kinetic reactions, respectively. Each contaminant and/or biological reaction must be represented by corresponding algebraic or ordinary differential equations depending upon the rate of the reaction. Since the reaction of one species depends upon the concentration of many other species, the final sets of equations are tightly coupled. For complex geochemical systems, consisting of many components and multidimensional transport, numerical solution of these coupled equations is challenging (Šimůnek and Valocchi 2002). As an alternative, more general models have recently been developed that more loosely couple transport and chemistry using a variety of sequential iterative or non-iterative operator-splitting approaches (e.g., Bell and Binning 2004; Jacques and Šimůnek 2005; Jacques et al. 2006). Models based on these various approaches are further discussed in Section 18.5.2.2.

18.3.6 Multiphase Flow and Transport

While the transport of contaminants in variably saturated media generally involves two phases (i.e., the liquid phase and soil gas, with advection in the gaseous phase often being neglected), many contamination problems also increasingly involve *nonaqueous phase liquids* (NAPLs) that are often only slightly miscible with water. Nonaqueous phase liquids may consist of single organic compounds such as many industrial solvents, or of a mixture of organic compounds such as gasoline and diesel fuel. Some of these compounds can be denser than water (commonly referred to as *dense nonaqueous phase liquids,* or DNAPLs) or lighter than water (*light nonaqueous phase liquids,* or LNAPLs). Their fate and dynamics in the subsurface is affected by a multitude of compound-specific flow and multicomponent transport processes, including interphase mass transfer and exchange (also with the solid phase).

Multiphase fluid flow models generally require flow equations for each fluid phase (water, air, NAPL). Two-phase air-water systems hence could be modeled also using separate equations for air and water. This shows that the standard Richards equation (Eq. (18.11)) is a simplification of a more complete multiphase (air-water) approach in that the air phase is assumed to have a negligible effect on variably-saturated flow, and that the air pressure varies only little in space and time. This assumption appears adequate for most variably-saturated flow problems. Similar assumptions, however, are generally not possible when NAPLs are present. Mathematical descriptions of multiphase flow and transport in general hence require separate flow equations for each of the three fluid phases, mass transport equations for all organic components (including those associated with the solid phase), and appropriate equations to account for interphase mass transfer processes. We refer readers to reviews by Abriola et al. (1999) and Rathfelder et al. (2000) for discussions of the complexities involved in modeling systems subject to multiphase flow, multicomponent transport and interphase mass transfer. A useful overview of a variety of experimental approaches for measuring the physical and hydraulic properties of multi-fluid systems is given by Lenhard et al. (2002).

18.4 Analytical Models

18.4.1 Analytical Approaches

Many *analytical solutions* have been derived in the past of the contaminant transport equations, and they are now widely used for analyzing contaminant transport during steady-state flow (Šimůnek 2005). Although a large number of analytical solutions also exist for the unsaturated flow equation, they generally can be applied only to relatively simple flow problems. The majority of applications for water flow in the vadose zone require a numerical solution of the Richards equation.

Analytical methods are representative of the classical mathematical approach for solving differential equations to produce an exact solution for a particular problem. Analytical models usually lead to an explicit equation for the concentration (or the pressure head, water content, or temperature) at a particular time and location. One hence can evaluate the concentration directly without time stepping typical of numerical methods. While exceptions exist (e.g., Liu et al. 2000), analytical solutions usually can be derived only for simplified transport systems involving linearized governing equations, homogeneous soils, simplified geometries of the transport domain, and constant or highly simplified initial and boundary conditions. Unfortunately, analytical solutions for more complex situations, such as for transient water flow or nonequilibrium contaminant transport with nonlinear reactions, are generally not available and/or cannot be derived, in which case numerical models must be adopted (Šimůnek 2005).

Analytical solutions are usually obtained by applying various transformations (e.g., Laplace, Fourier or other transforms) to the governing equations, invoking a separation of variables, and/or using the Green's function approach (e.g., Leij et al. 2000).

18.4.2 Existing Models

18.4.2.1 One-Dimensional Models

Some of the more popular one-dimensional analytical transport models have been CFITM (Van Genuchten 1980b), CFITIM (Van Genuchten 1981), CXTFIT (Parker and Van Genuchten 1984), and CXTFIT2 (Toride et al. 1995). While CFITM considers only one-dimensional equilibrium transport in both finite and semi-infinite domains, CFITIM additionally considers physical and chemical nonequilibrium transport (i.e., the two-region mobile-immobile model for physical nonequilibrium and the two-site sorption model for chemical nonequilibrium). CXTFIT expanded the capabilities of CFITIM by considering more general initial and boundary conditions, as well as degradation processes. CXTFIT2 (Toride et al. 1995), an updated version of CXTFIT, solves both direct and inverse problems for three different one-dimensional transport models:

- the conventional advection-dispersion equation, ADE;
- the chemical and physical nonequilibrium ADEs;
- a stochastic stream tube model based upon the local-scale equilibrium or nonequilibrium ADE.

These three types of models all consider linear adsorption, and include zero- and first-order decay/source terms.

18.4.2.2 Multi-Dimensional Models

Some of the more popular multi-dimensional analytical transport models have been AT123D (Yeh 1981), 3DADE (Leij and Bradford 1994), N3DADE (Leij and Toride 1997), and MYGRT (Ungs et al. 1998). These programs provide analytical solutions to transport problems in two- and three-dimensional domains. 3DADE also includes parameter estimation capabilities.

A large number of analytical models for one-, two-, and three-dimensional contaminant transport problems were incorporated into the public domain software package STANMOD (STudio of ANalytical MODels) (Šimůnek et al. 1999a) (http://www.hydrus2d.com). This Windows-based computer software package includes not only programs for equilibrium advective-dispersive transport such as the CFITM code of Van Genuchten (1980b) for one-dimensional transport and 3DADE (Leij and Bradford 1994) for three-dimensional problems, but also programs for more complex problems. For example, STANMOD also incorporates the CFITIM (Van Genuchten 1981) and N3DADE (Leij and Toride 1997) programs for nonequilibrium transport (i.e., the two-region mobile-immobile model for physical nonequilibrium and the two-site sorption model for chemical nonequilibrium) in one and multiple dimensions, respectively. A more recent version of STANMOD includes additionally the screening model of Jury et al. (1983a) for transport and volatilization of soil-applied organic contaminants.

18.5 Numerical Models

18.5.1 Numerical Approaches

Although analytical and semi-analytical solutions are still popularly used for solving many relatively simple problems, the ever-increasing power of personal computers and the development of more accurate and stable numerical solution techniques have led to the much wider use of *numerical models* over the past ten years. Numerical methods in general are superior to analytical methods in terms of their ability to solve much more realistic problems (Šimůnek 2005). They allow users to design complicated geometries that reflect complex natural pedological and hydrological conditions, control parameters in space and time, prescribe more realistic initial and boundary conditions, and permit the implementation of nonlinear constitutive relationships (Šimůnek and Van Genuchten 2006). Numerical methods subdivide the time and spatial coordinates into smaller pieces, such as finite differences, finite elements, and/or finite volumes, and reformulate the continuous form of governing partial differential equations in terms of a system of algebraic equations. In order to obtain solutions at prescribed times, numerical methods generally require intermediate simulations (time-stepping) between the initial condition and the points in time for which the solution is needed.

Reviews of the history of development of various numerical techniques used in vadose zone flow and contaminant transport models are given by Van Genuchten and Šimůnek (1996) and Šimůnek (2005).

18.5.1.1 Finite Differences

Finite difference methods are generally very intuitive and relatively easy to implement. Time and space are divided into small increments Δt and Δz (or Δx and Δz) (Fig. 18.16). Temporal and spatial derivatives in the governing equations are then replaced with finite differences (formally using Taylor series expansions). For example, the standard advection-dispersion equation for steady-state water flow (Wang and Anderson 1982) given by

$$\frac{\partial c}{\partial t} = -\frac{\partial J}{\partial z} = D\frac{\partial^2 c}{\partial z^2} - v\frac{\partial c}{\partial z} \tag{18.58}$$

can be approximated as follows using an explicit (forward-in-time) finite difference scheme:

$$\frac{c_i^{j+1} - c_i^j}{\Delta t} = -\frac{J_{i+1/2}^j - J_{i-1/2}^j}{\Delta z} = D\frac{c_{i+1}^j - 2c_i^j + c_{i-1}^j}{(\Delta z)^2} - v\frac{c_{i+1}^j - c_{i-1}^j}{2\Delta z} \tag{18.59}$$

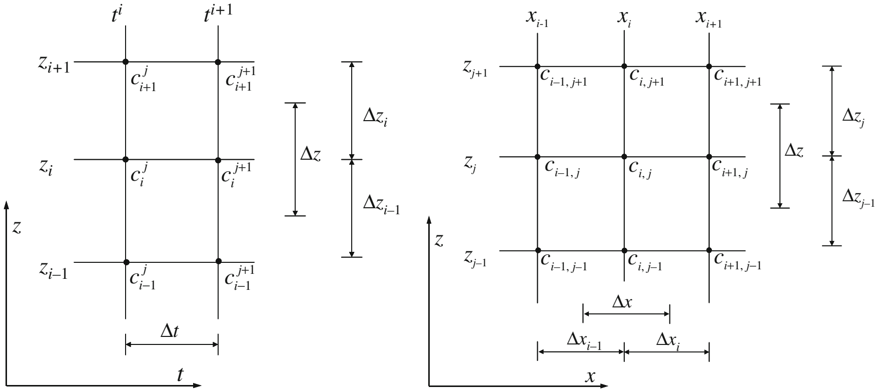


Fig. 18.16 Examples of the spatial and temporal difference discretization of a one-dimensional problem (*left*), and the finite difference discretization of a two-dimensional domain (*right*) (after Šimůnek and Van Genuchten (2006))

where subscripts refer to spatial discretization and superscript to temporal discretization (e.g., j and j+1 are for the previous and actual time levels, respectively; see Fig. 18.16), Δt is the time step, and Δz is the spatial step (assumed to be constant). Notice that this equation contains only one unknown variable (i.e., the concentration c_i^{j+1} at the new time level), which hence can be evaluated directly (explicitly) by solving the equation.

By comparison, a fully implicit (backward-in-time) finite difference scheme can be written as follows

$$\frac{c_i^{j+1}-c_i^j}{\Delta t}=-\frac{J_{i+1/2}^{j+1}-J_{i-1/2}^{j+1}}{\Delta z}=D\frac{c_{i+1}^{j+1}-2c_i^{j+1}+c_{i-1}^{j+1}}{(\Delta z)^2}-v\frac{c_{i+1}^{j+1}-c_{i-1}^{j+1}}{2\Delta z} \tag{18.60}$$

and an implicit (weighted) finite difference scheme as:

$$\begin{aligned}\frac{c_i^{j+1}-c_i^j}{\Delta t}=&D\frac{\varepsilon\left(c_{i+1}^{j+1}-2c_i^{j+1}+c_{i-1}^{j+1}\right)+(1-\varepsilon)\left(c_{i+1}^{j}-2c_i^{j}+c_{i-1}^{j}\right)}{(\Delta z)^2}\\&-v\frac{\varepsilon\left(c_{i+1}^{j+1}-c_{i-1}^{j+1}\right)+(1-\varepsilon)\left(c_{i+1}^{j}-c_{i-1}^{j}\right)}{2\Delta z}\end{aligned} \tag{18.61}$$

where ε is a temporal weighting coefficient. Different finite difference schemes result depending upon the value of ε, i.e., an explicit scheme when $\varepsilon = 0$, a Crank-Nicholson time-centered scheme when $\varepsilon = 0.5$, and a fully implicit scheme when ε=1.

18.5.1.2 Finite Elements

Finite element methods can be implemented in very much the same way as finite differences for one-, two-, and three-dimensional problems. A major advantage of the finite elements is that they are much easier used to discretize complex two- and three-dimensional transport domains (Fig. 18.17). As an example, Fig. 18.17 shows triangular unstructured finite element grids for a regular rectangular and an irregular domain as generated with the automated MeshGen2D mesh generator of HYDRUS-2D (Šimůnek et al. 1999b). Notice that even though the figure on the right (Fig. 18.17) has an irregular soil surface, as well as a tile drain within the transport domain, MeshGen2D could easily discretize/accommodate this transport domain using an unstructured triangular finite element mesh.

18.5.2 Existing Models

18.5.2.1 Single-Species Solute Transport Models

A large number of numerical models are now available for evaluating variably-saturated water flow and contaminant transport processes in the subsurface. Some of these models are in the public domain, such as MACRO (Jarvis et al. 1994), SWAP (van Dam et al. 1997), UNSATH (Fayer 2000), VS2DI (Healy 1990), and

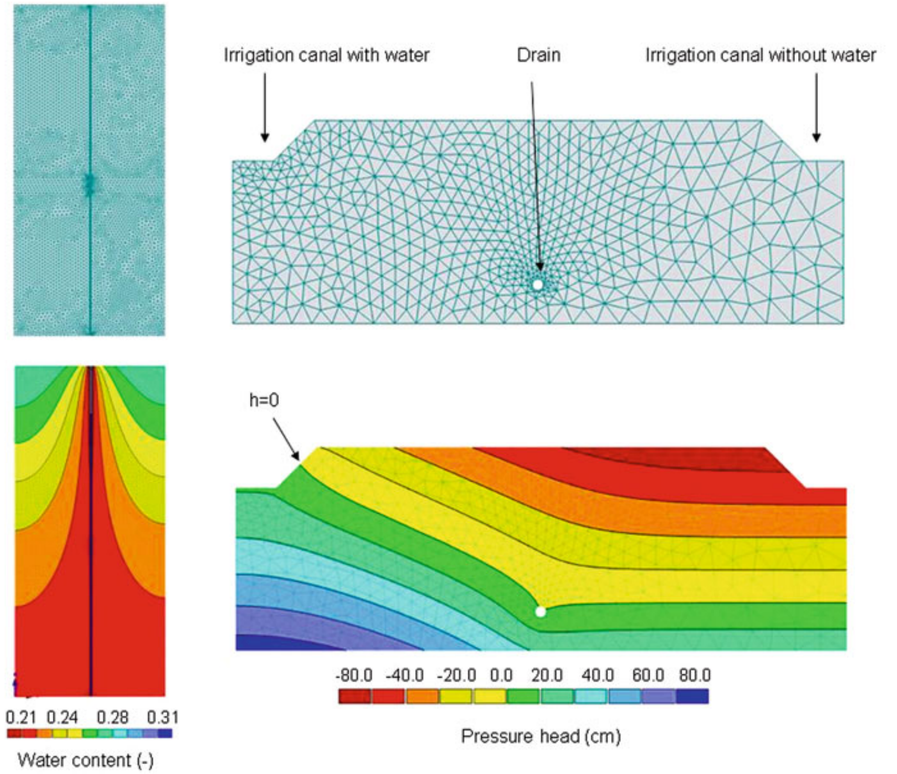


Fig. 18.17 Examples of triangular finite element grids for regular (*left*) and irregular (*right*) two-dimensional transport domains. The problem on the left displays a 1-m wide and 2-m high soil profile with a 0.01-cm wide macropore in the middle and a non-uniform steady-state water content distribution. The project on the right shows a drain in between two furrows (one dry, one filled with water) and the resulting steady-state pressure head distribution (groundwater table corresponds to $h = 0$)

HYDRUS-1D (Šimůnek and Van Genuchten 2008; Šimůnek et al. 1998a, 2005), while others are in the commercial domain, such as HYDRUS-2D (Šimůnek et al. 1999b), HYDRUS (2D/3D) (Šimůnek et al. 2006c), and MODFLOW-SURFACT (HydroGeoLogic 1996). The models vary widely in terms of their complexity, sophistication, and ease of use. Although some models are still being run under the DOS operating system, with associated difficulties of preparing input files and interpreting tabulated outputs, many others, especially those in the commercial domain, are supported by sophisticated graphics-based interfaces that greatly simplify their use (Šimůnek et al. 1998a, 1999b). Several studies have recently reviewed and compared various numerical models for vadose zone applications (e.g., MDH Engineered Solutions Corp. 2003; Scanlon et al. 2002; Vanderborght et al. 2005; Wilson et al. 1999). These studies typically compared the precision, speed, and ease of use of the codes involved.

While earlier models solved the governing flow and transport equations for relatively simplified system-independent boundary conditions (i.e., specified pressure heads or fluxes, and free drainage), the more recent models can cope with much more complex system-dependent boundary conditions evaluating surface flow and energy balances and accounting for the simultaneous movement of water, vapor, and heat. Examples are DAISY (Hansen et al. 1990), TOUGH2 (Pruess 1991), SHAW (Flerchinger et al. 1996), SWAP (Van Dam et al. 1997), HYDRUS-1D (Šimůnek et al. 1998a, 2005), UNSATH (Fayer 2000), and COUP (Jansson and Karlberg 2001). Several models now also account for the extremely nonlinear processes associated with the freezing and thawing cycle (e.g., DAISY, SHAW, and COUP).

Contaminant transport models have also become more sophisticated in terms of the type and complexity of processes that can be simulated. Transport models are no longer being limited to contaminants undergoing relatively simple chemical reactions such as linear sorption and first-order decay, but now consider also a variety of nonlinear sorption and exchange processes, physical and chemical nonequilibrium transport, volatilization, gas diffusion, colloid attachment/ detachment, decay chain reactions, and many other processes (e.g., the HYDRUS-1D, –2D, and (2D/3D) codes of Šimůnek et al. (1999b, 2005, 2006c), or MODFLOW-SURFACT of HydroGeoLogic, Inc. (1996)). For example, the general formulation of the transport equations in the HYDRUS codes permit simulations of non-adsorbing or linearly sorbing contaminants, in addition to a variety of other contaminants, such a viruses (Schijven and Šimůnek 2002), colloids (Bradford et al. 2002), cadmium (Seuntjens et al. 2001), and hormones (Casey et al. 2003, 2004), or contaminants involved in the sequential biodegradation of chlorinated aliphatic hydrocarbons (Casey and Šimůnek 2001; Schaerlaekens et al. 1999).

Much effort has been directed also toward improving models for purposes of simulating nonequilibrium and/or preferential flow. Examples are the TOUGH codes (Pruess 1991, 2004), MACRO (Jarvis et al. 1994), and HYDRUS-1D (Šimůnek and Van Genuchten 2008). These models typically assume the presence of dual-porosity and dual-permeability regions, with different fluxes possible in the two regions. Example applications of these dual-porosity and dual-permeability models are given by Mallants et al. (1997), Šimůnek et al. (2001), Haws et al. (2005), Köhne et al. (2004, 2006, 2009a, b), and Pot et al. (2005), among many others.

As an example of available vadose zone flow and transport models, we briefly discuss here the HYDRUS software packages of Šimůnek et al. (1999b, 2005, 2008).

The HYDRUS Software Packages

HYDRUS-1D (Šimůnek et al. 2005), HYDRUS-2D (Šimůnek et al. 1999b), and HYDRUS (2D/3D) (Šimůnek et al. 2006c) are software packages (http://www.pc-progress.com/en/Default.aspx) that simulate the one- and two-dimensional movement of water, heat, and multiple contaminants in variably saturated porous media, respectively. Both programs use finite elements to numerically solve the Richards equation for saturated-unsaturated water flow and Fickian-based advection-dispersion equations for both heat and contaminant transport. The unsaturated soil hydraulic properties can be described using Van Genuchten (1980a), Brooks and Corey (1964), Kosugi (1996), and Durner (1994) type analytical functions, or modified Van Genuchten type functions that produce a better description of the hydraulic properties near saturation.

The HYDRUS-1D software package additionally includes modules for simulating carbon dioxide and major ion contaminant movement (Šimůnek et al. 1996; Šimůnek and Suarez 1993). Also included is a small catalog of unsaturated soil hydraulic properties (Carsel and Parish 1988), as well as pedotransfer functions based on neural network predictions (Schaap et al. 2001).

18.5.2.2 Biogeochemical Transport Models

Significant efforts have been made also in coupling physical flow and transport models with biogeochemical models to simulate increasingly more complex reactions, such as surface complexation, precipitation/dissolution, cation exchange, and/or (micro)biological reactions. Reviews of the development of hydrogeochemical transport models involving reactive multiple components are given by Mangold and Tsang (1991), Lichtner (1996), Steefel and MacQuarrie (1996), Šimůnek and Valocchi (2002), and Bell and Binning (2004). Most modeling efforts involving multicomponent transport have thus far focused on the saturated zone, where changes in the flow velocity, temperature and pH are often much more gradual and hence less important than in the unsaturated zone. Consequently, most multicomponent transport models assumed one- or two-dimensional steady-state saturated water flow with a fixed value of the flow velocity, temperature and *p*H. Several multicomponent transport models have been published also for variably-saturated flow problems. These include DYNAMIX (Liu and Narasimhan 1989), HYDROGEOCHEM (Yeh and Tripathi 1990), TOUGH-REACT (Pruess 1991), UNSATCHEM (Šimůnek and Suarez 1994; Šimůnek et al. 1996, 1997), FEHM (Zyvoloski et al. 1997), MULTIFLO (Lichtner and Seth 1996), OS3D/GIMRT (Steefel and Yabusaki 1996), HYDROBIOGEOCHEM (Yeh et al. 1998), FLOTRAN (Lichtner 2000), MIN3P (Mayer et al. 2002), HP1 (Jacques and Šimůnek 2005; Jacques et al. 2002, 2008a, 2008b), and HYDRUS-1D (Šimůnek et al. 2005).

Geochemical models can be divided into two major groups: those with specific chemistry and those characterized by more general chemistry (Šimůnek and Valocchi 2002). Models with specific chemistry are limited in the number of species they can handle, while their application is restricted to problems having a prescribed chemical system. They are, however, much easier to use and computationally can be much more efficient than general models. Typical examples of models with specified chemistry are those simulating the transport of major ions, such as LEACHM (Wagenet and Hutson 1987), UNSATCHEM (Šimůnek and Suarez 1994; Šimůnek et al. 1996), and HYDRUS-1D (Šimůnek et al. 2005). Models with generalized chemistry (DYNAMIX, HYDROGEOCHEM, MULTIFLO, FLOTRAN, OS3D/GIMRT, and HP1, all referenced above) provide users with much more freedom in designing a particular chemical system; possible applications of these models are also much wider.

HP1

HYDRUS-1D was coupled with the PHREEQC geochemical code (Parkhurst and Appelo 1999) to create a new comprehensive simulation tool, HP1 (acronym for HYDRUS1D-PHREEQC) (Jacques and Šimůnek 2005; Jacques et al. 2003, 2008a, 2008b). The combined code contains modules simulating (1) transient water flow in variably-saturated media, (2) the transport of multiple components, (3) mixed equilibrium/kinetic biogeochemical reactions, and (4) heat transport. HP1 is a significant expansion of the individual HYDRUS-1D and PHREEQC programs by preserving most of their original features and capabilities. The code still uses the Richards equation for simulating variably-saturated water flow and advection-dispersion type equations for heat and contaminant transport. However, the program can now simulate also a broad range of low-temperature biogeochemical reactions in water, the vadose zone and in ground water systems, including interactions with minerals, gases, exchangers, and sorption surfaces, based on thermodynamic equilibrium, kinetics, or mixed equilibrium-kinetic reactions.

Jacques et al. (2003, 2008a, 2008b) and Jacques and Šimůnek (2005) demonstrated the versatility of the HP1 model on several examples such as:

- the transport of heavy metals (Zn^{2+}, Pb^{2+}, and Cd^{2+}) subject to multiple cation exchange reactions;
- transport with mineral dissolution of amorphous SiO_2 and gibbsite ($Al(OH)_3$);
- heavy metal transport in a medium with a pH-dependent cation exchange complex;
- infiltration of a hyperalkaline solution in a clay sample (this example considers kinetic precipitation-dissolution of kaolinite, illite, quartz, calcite, dolomite, gypsum, hydrotalcite, and sepiolite);
- long-term transient flow and transport of major cations (Na^+, K^+, Ca^{2+}, and Mg^{2+}) and heavy metals (Cd^{2+}, Zn^{2+}, and Pb^{2+}) in a soil profile
- cadmium leaching in acid sandy soils;
- radionuclide transport following phosphorus fertilization (U and its aqueous complexes);
- the fate and subsurface transport of explosives (TNT and its daughter products 2ADNT, 4ADNT, and TAT) (Šimůnek et al. 2006b).

To illustrate the capabilities of HP1 and the importance of coupling in an integrated manner all relevant physical and contaminant transport processes when analyzing the fate and transport of contaminants in variably saturated field soils, the next section discusses the leaching of cadmium from a podsol soil under transient boundary conditions.

Applications to Unsaturated Flow and Geochemical Transport Modeling

In this example (Jacques et al. 2008a) we discuss a hypothetical HP1 application involving the transport of major cations and heavy metals in a soil during transient flow over a period of 30 years. Results will show that variations in water contents and water fluxes can significantly influence the speciation, and thus the mobility and availability, of elements. Decreasing water contents near the soil surface, furthermore, may lower the pH of the soil liquid phase and produce new cation exchange equilibrium conditions. The upward transport of Cl during summer due to increased evapotranspiration, and subsequent accumulation of Cl near the soil surface, can cause an increase in the total aqueous cadmium concentration because of the formation of cadmium-Cl complexes.

The HP1 multi-component transport simulator is used in this section to simulate the leaching of cadmium and zinc in a dry Spodosol in a sandy region of Northern Belgium using in-situ measured cadmium and zinc (and additional elements) concentration profiles. Soils in the region were contaminated by atmospheric deposition of cadmium and zinc from non-ferrous industry (Seuntjens 2000). Water flow and the transport of major cations (Na, K, Mg, and Ca), heavy metals (cadmium and zinc), anions (Cl, Br) and Al were simulated for a 1-m deep multi-layered podsol soil profile subject to atmospheric boundary conditions for a period of 30 years. The main focus is on how processes affecting water contents and water fluxes also influence the geochemical conditions in the soil. Specifically, the effect of cycles of evaporation and infiltration on pH and cadmium speciation will be discussed.

Interactions between major cations and heavy metals with the soil solid phase were simulated by means of cation exchange processes assuming local equilibrium on a single type of exchange sites. Voegelin (2001) earlier showed that this approach adequately describes various features of cadmium transport experiments. Although sorption on specific sites with a high affinity for cadmium may also occur in soils (Selim et al. 1992), this type of binding is unlikely in acid sandy soils (Voegelin et al. 2001). Exchange parameters were calibrated using concentrations measured in drainage water from a steady-state flow experiment on large, undisturbed soil lysimeters (1 m long and 0.8 m diameter; Seuntjens et al. 2001). The initial composition of the cation exchange site was measured for each soil horizon (Jacques et al. 2008a). While the same log(K) parameters were used for all soil horizons, the size of the cation exchange complex was assumed to vary between horizons.

Daily values of precipitation (P) and potential evapotranspiration (Ep) during the 30-year simulation period were imposed as climatic boundary condition, from which the daily actual evaporation rate was calculated. A subset of the P-E_p data is shown in Fig. 18.18.

Figure 18.19 shows time series between 1972 and 1982 for the water content, pH and total Cl and Cd concentrations in the liquid phase at two depths. The results illustrate the relation between water flow and geochemical conditions in the soil. The alternation between precipitation (wet conditions) and evaporation (dry conditions) as dictated by the atmospheric conditions clearly affected the dynamics of the water content, with upward water flow during dry periods. The flow dynamics in turn significantly influenced the geochemistry near the soil surface. As illustrated, the most mobile elements (anions such as Cl^- and monovalent cations such as Na^+) move upwards during the evaporation periods, thus causing these ions to accumulate near the soil surface. The decrease in water content near the soil surface due to evaporation resulted in higher concentrations and a lower pH. PHREEQC

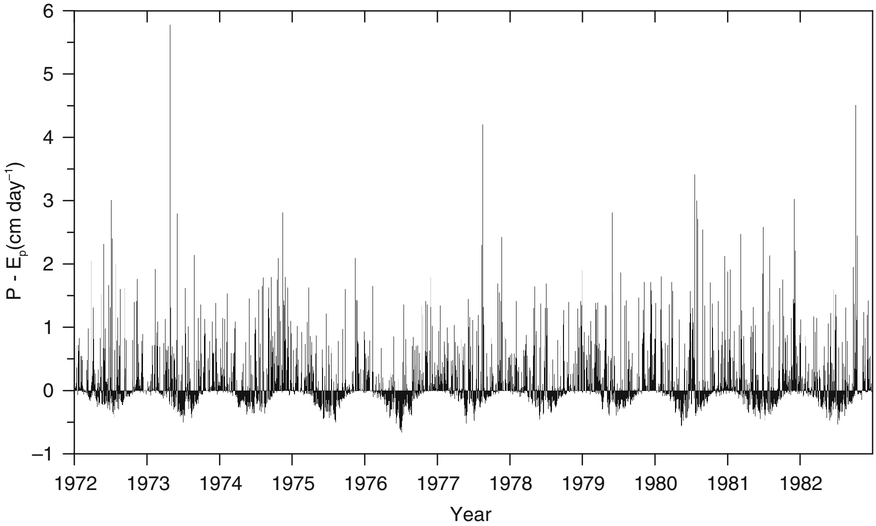


Fig. 18.18 Potential precipitation surplus during years 1972–1982

calculations showed that removing only water from aqueous solutions results in a pH decrease (for a similar geochemical system, i.e., only cation exchange and equilibrium with gibbsite). Another factor intensifying the pH decrease was the greater mobility of anions (Cl^- and Br^-) compared to cations. Thus, the physical factors of having a decreasing water content and upward flow of water and contaminants caused the pH to decrease when $P - E_p$ is negative. It should be noted that the *p*H in reality is also affected by other geochemical or biological processes not included in the invoked conceptual model. For example, soil carbon dioxide concentrations that usually change in response to biological activity and moisture status of the soil can also affect soil *p*H (Šimůnek and Suarez 1993).

Although upward flow during the summer had almost no effect on the total amount of heavy metals in the surface horizon due to the low mobility of these elements, the aqueous concentrations of the metals did vary significantly during the season. Several factors contributed to this. First, because of lower water contents, the concentration of all aqueous species increased during summer periods. Changes in aqueous concentrations in turn caused changes in the cation exchange equilibrium, thereby promoting monovalent cations to sorb onto the cation exchange complex and bivalent cations to desorb into solution. This explains also the difference between $T_{Na,aq}$ and $T_{Ca,aq}$ during the summer near the soil surface. The aqueous concentration of Na was controlled more by the cation exchange complex than that of Ca due to preferred adsorption of Na during dry soil conditions. This process is further amplified by the increased supply of monovalent cations due to upward flow of water during summer, leading to relatively more sorption of the monovalent cations and higher concentrations of divalent cations and heavy metals in the liquid phase. The complexation of cadmium with Cl contributed also to the

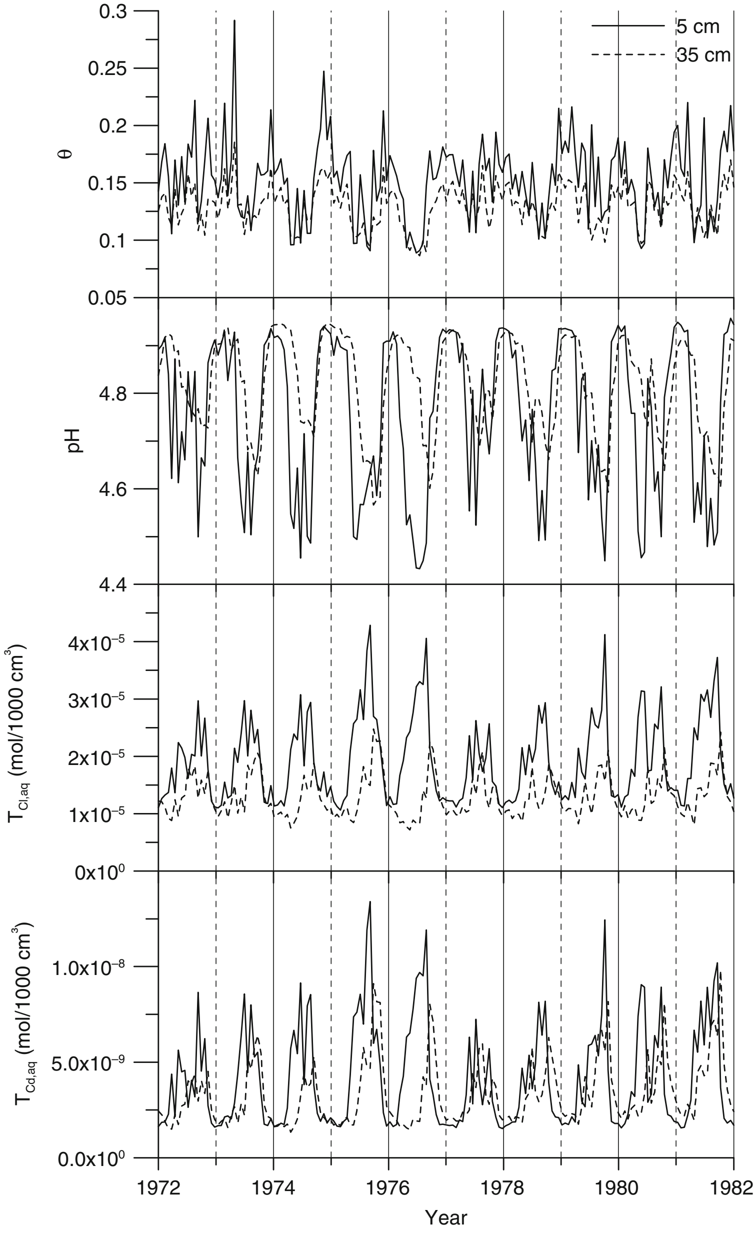


Fig. 18.19 Time series of water content (θ), pH, $T_{cl,aq}$ and $T_{Cd,aq}$ (mol/1000 cm^3 soil) between 1972 and 1982 at 5 cm (A-horizon) and 35 cm (Bh1-horizon)

increase in $T_{Cd,aq}$. At a depth of 3 cm, about 0.65 and up to 3.5 percent of the aqueous cadmium was in the form of $CdCl^+$ during winter and summer, respectively. Other cadmium complexes of the form $CdCl_n^{(2-n)}$ were present in much smaller concentrations. Similar to the pH changes, changing water contents and upward fluxes both affected the amount of cadmium in the liquid phase.

Our results above indicate that atmospheric boundary conditions can have a significant effect on the amount and transport of cadmium in a soil profile, and on its bioavailability since uptake processes by plants and soil micro-organisms are often concentration-dependent. Passive root uptake of contaminants together with water increases with increasing contaminant concentrations. Similarly, active uptake as described with Monod or Michealis-Menten kinetics will increase with increasing contaminant concentrations. Moreover, the high heavy metal concentrations occurred during the summer months with the highest (micro)biological activity. In addition, cadmium speciation may also play a role in uptake. For example, Smolders and McLaughlin (1996) observed more cadmium uptake by chard (*Beta vulgaris* var. *cicla*) when Cl concentrations increased while the Cd^{2+} activity was kept constant, likely due to phytoavailability of $CdCl^+$ and other $CdCl_n{}^{2-n}$-species.

The example illustrates that simulators such as HP1 are potentially attractive tools for studying reactive transport processes in the vadose zone during transient variably-saturated flow. Geochemical conditions are an important factor since they determine the speciation (both in the liquid phase and on the solid phase) of the elements involved, and thus their mobility and bioavailability. Small variations in prevailing geochemical conditions may alter significantly the speciation and mobility of heavy metals or other constituents. Soil systems, moreover, are subject to large transient variations since they are open to the atmosphere. Hence, changes in the composition of rain water or atmospheric deposition, among other external factors, can materially alter the geochemical conditions in a soil profile.

For the time being, only a limited number of test cases and experimental data exist for the unsaturated zone to assess the full capabilities of a reactive transport code (see also Davis et al. 2004). One elaborate recent application is a study by Gonçalves et al. (2006), who successfully applied the chemistry-specific major ion geochemistry module of HYDRUS-1D (version 3.0, Šimůnek et al. 2005) to a 4-year experimental data set involving water flow and contaminant transport in lysimeters irrigated with waters of different quality and subjected to atmospheric conditions. Additional studies of this type should provide more credibility to the use of coupled hydrogeochemical models for addressing flow and reactive transport problems in the vadose zone.

18.6 Concluding Remarks

This chapter demonstrates the abundance of models and modeling approaches that are currently available for simulating variably-saturated water flow and contaminant transport at various levels of approximation and for different applications. Models range from relatively simple analytical approaches for analyzing contaminant transport problems during one-dimensional steady-state flow, to sophisticated numerical models for addressing multi-dimensional variably-saturated flow and contaminant transport problems at the field scale.

One may expect that unsaturated zone flow and transport models will be used increasingly as tools for higher tier Groundwater-related Risk Assessments (see Chapter 17 by Swartjes and Grima, this book). Moreover, these models may support the development of cost-effective, yet technically sound strategies for resource management, contamination remediation, and/or prevention. Improved understanding of the key underlying processes, continued advances in numerical methods, and the introduction of more and more powerful computers make such simulations increasingly practical for many field-scale problems. Models can be helpful tools also for designing, testing and implementing soil, water and crop management practices that minimize soil and water contamination. Models are equally needed for designing or remediating industrial waste disposal sites and landfills, for predicting contaminant transport from mining wastes, or for long-term stewardship of nuclear waste repositories. A major challenge is to make the models as realistic as possible for such complex applications. Yet another challenge is to obtain input data appropriate for the time and spatial scale under consideration, accounting for spatial variability and possibly even time-dependent parameters. Several approaches were discussed that permit generation of input data from other more basic (soil) data, such as pedotransfer functions. These efforts must continue to also include (bio)chemical parameters.

Continued progress in subsurface flow and transport modeling requires equal advances in both numerical techniques as well as the underlying science. Addressing preferential flow phenomena, and the related problems of subsurface heterogeneity, including the stochastic nature of boundary conditions (precipitation and/or evapotranspiration), will continue to pose formidable challenges. The same is true for improving multicomponent geochemical transport modeling for the vadose zone. For example, numerical algorithms and databases for multicomponent transport models must be extended to higher temperatures and ionic strengths, complex contaminant mixtures (including especially mixed organic and inorganic wastes), multiphase flow, redox disequilibria for low-temperature systems, and coupled physico-chemical systems to account for possible changes in the water retention and hydraulic conductivity functions. Better integration is also needed between variably-saturated zone and existing larger-scale surface numerical models, which in turn requires further research on such issues as spatial and temporal scaling of hydrological, chemical and biological processes and properties, linking constitutive (soil hydraulic) relationships to measurements scales, preferential flow, and issues of parameter and model uncertainty.

Many scientific questions related to colloid and colloid-facilitated transport are also still largely unresolved. This is an area of research where our understanding lags far behind current numerical capabilities. Much work is needed to better understand the processes of filtration, straining, size exclusion, colloid-colloid interactions, mobilization of colloids and microorganisms; accumulation at air-water interfaces, interactions between microorganisms and contaminants (including biodegradation), the effects of both physical factors (water content, flow velocity, textural interfaces) and chemical processes (ionic strength, solution composition, *p*H) on colloid retention and mobilization, and modeling colloid-facilitated transport during conditions of transient flow.

Also, to the best of our knowledge, no models are currentely available that consider all of the many processes simultaneously and in their full complexity, including their mutual interactions. That is, no models exist that consider transient preferential flow and transport in structured soils or fractured rocks, while simultaneously considering complex biogeochemical reactions between contaminants, organic and inorganic colloids and/or organic complexes, and solid and air phases of a soil, including widely varying rates of these various reactions. Further integration of the different types of numerical models is needed to address practical problems of contaminant transport (trace elements, radionuclides, organic contaminants) in complex vadose zone environments.

References

Abriola LM, Pennell KD, Weber WJ Jr, Lang JR, Wilkins MD (1999) Persistence and interphase mass transfer of organic contaminants in the unsaturated zone: experimental observations and mathematical modeling. In: Parlange JY, Hopmans JW (eds) Vadose zone hydrology: cutting across disciplines. Oxford University, New York, pp 210–234

Ahuja LR, Hebson C (1992) Root zone water quality model. GPSR Tech Rep No 2, USDA, ARS, Fort Collins, CO

Allen RG (2001) REF-ET: reference evapotranspiration calculation software for FAO and ASCE standardized equations, version 2.0 for windows. University of Idaho, USA

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: guidelines for computing crop water requirements, Irrigation and Drainage Paper, no. 56, Food and Agriculture Organization of the United Nations, Rome

Anderson MP (1984) Movement of contaminants in groundwater: groundwater transport-advection and dispersion. In: Groundwater contamination. Studies in geophysics. National Academy, Washington, DC, pp 37–45

Babalola O (1978) Spatial variability of soil water properties in tropical soils of Nigeria. Soil Sci 126:269–279

Bear J (1972) Dynamics of fluids in porous media. Elsevier Science, NY

Bell LSJ, Binning PJ (2004) A split operator approach to reactive transport with the forward particle tracking Eulerian-Lagrangian localized adjoint method. Adv Water Resour 27: 323–334

Biggar JW, Nielsen DR (1967) Miscible displacement and leaching phenomena. Agronomy 11:254–274

Biggar JW, Nielsen DR (1976) Spatial variability of the leaching characteristics of a field soil. Water Resour Res 12:78–84

Bouma J (1989) Using soil survey data for quantitative land evaluation. Adv Soil Sci 9: 177–213

Bradford SA, Šimůnek J, Bettahar M, Van Genuchten MTh, Yates SR (2003) Modeling colloid attachment, straining, and exclusion in saturated porous media. Environ Sci Technol 37: 2242–2250

Bradford SA, Yates SR, Bettahar M, Šimůnek J (2002) Physical factors affecting the transport and fate of colloids in saturated porous media. Water Resour Res 38(12):1327. doi:10.1029/2002WR001340

Brooks RH, Corey AT (1964) Hydraulic properties of porous media. Hydrol Paper No. 3, Colorado State University, Fort Collins, CO

Brusseau ML (1999) Nonideal transport of reactive solutes in porous media; Cutting across history and disciplines. In: Parlange MB, Hopmans JW (eds) Vadose zone hydrology: cutting across disciplines. Oxford University, New York, NY, pp 130–154

Burdine NT (1953) Relative permeability calculations from pore-size distribution data. Petr Trans Am Inst Mining Metall Eng 198:71–77

Cameron DR, Nyborg M, Toogood JA, Laverty DH (1971) Accuracy of field sampling for soil tests. Can J Soil Sci 51:165–175

Carsel RF, Parrish RS (1988) Developing joint probability distributions of soil water retention characteristics. Water Resour Res 24:755–769

Casey FXM, Šimůnek J (2001) Inverse analyses of the transport of chlorinated hydrocarbons subject to sequential transformation reactions. J Environ Qual 30(4):1354–1360

Casey FXM, Larsen GL, Hakk H, Šimůnek J (2003) Fate and transport of 17β-Estradiol in soil-water systems. Environ Sci Technol 37(11):2400–2409

Casey FXM, Larsen GL, Hakk H, Šimůnek J (2004) Fate and transport of testosterone in agriculturally significant soils. Environ Sci Technol 38(3):790–798

Cassel DK (1983) Spatial and temporal variability of soil physical properties following tillage of Norfolk loamy sand. Soil Sci Soc Am J 47:196–201

Cassel DK, Bauer, A (1975) Spatial variability in soils below depth of tillage. Bulk density and fifteen atmosphere percentage. Soil Sci Soc Am J 39:247–250

Cosby BJ, Hornberger GM, Clapp RB, Ginn TR (1984) A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour Res 6: 682–690

Courtin P, Feller MC, Klinka K (1983) Lateral variability in some properties of disturbed forest soils in SW British Columbia. Can J Soil Sci 63:529–539

Dagan G (1989) Flow and transport in porous formations. Springer, Berlin

Dane JH, Hopmans JW (2002) Water retention and storage. In: Dane JH, Topp GC (eds) Methods of soil analysis, part 1, Physical methods, chapter 3.3, 3rd edn. SSSA, Madison, WI, pp 671–720

Dane JH, Topp GC (2002) Methods of soil analysis, part 1, physical methods, 3rd edn. SSSA, Madison, WI

Davis JA, Yabusaki SB, Steefel CI, Zachara JM, Curtis GP, Redden GB, Criscenti LJ, Honeyman BD (2004) Assessing conceptual models for subsurface reactive transport of inorganic contaminants. EOS 85:449, 455

de Jonge LW, Kjaergaard C, Moldrup P (2004) Colloids and colloid-facilitated transport of contaminants in soils: an introduction. Vadose Zone J 3:321–325

Delhomme JP (1979) Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. Water Resour Res 15:269–280

Durner W (1994) Hydraulic conductivity estimation for soils with heterogeneous pore structure, Water Resour Res 32(9):211–183

El Abd H (1984) Spatial variability of the pesticide distribution coefficient. PhD thesis, University of California, Riverside, 100 pp

Fayer MJ (2000) UNSAT-H Version 3.0: unsaturated soil water and heat flow model. Theory, user manual, and examples. Pacific Northwest National Laboratory 13249

Feddes RA, Kowalik PJ, Zaradny H (1978) Simulation of field water use and crop yield. Wiley, New York, NY

Fetter CW (1999) Contaminant hydrology. Upper Saddle River, Prentice Hall, 458 pp

Flerchinger GN, Hanson CL, Wight JR (1996) Modeling evapotranspiration and surface energy budgets across a watershed. Water Resour Res 32:2539–2548

Flury M, Fluhler H, Jury W, Leuenberger J (1994) Susceptibility of soils to preferential flow of water: a field study. Water Resour Res 30:1945–1954

Flury M, Qiu H (2008) Modeling colloid-facilitated contaminant transport in the Vadose zone. Vadose Zone J 7:682–697

Freundlich H (1909) Kapillarchemie; eine Darstellung der Chemie der Kolloide und verwandter Gebiete, Akademische Verlagsgellschaft, Leipzig, Germany

Gajem YM, Warrick AW, Myers DE (1981) Spatial dependence of physical properties of a typic torrifluvent soil. Soil Sci Soc Am J 45:709–715

Gelhar LW, Mantaglou A, Welty C, Rehfeldt KR (1985) A review of field scale physical solute transport processes in unsaturated and saturated porous media. EPRI Topicl Report EA-4190, Electric Power Research Institute, Palo Alto, CA

Gerke HH, Van Genuchten MTh (1993) A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. Water Resour Res 29:305–319

Gerke HH, Van Genuchten MTh (1996) Macroscopic representation of structural geometry for simulating water and solute movement in dual-porosity media. Adv Water Resour 19:343–357

Glotfelty DE, Schomburg CJ (1989) Volatilization of pesticides. In: Sawhney BL, Brown K (eds) Reactions and movement of organic chemicals in soils, special publ. no. 18. Soil Science Society of America, Madison, WI, pp 181–208

Gonçalves MC, Šimůnek J., Ramos TB, Martins JC, Neves MJ, Pires FP (2006) Multicomponent solute transport in soil lysimeters irrigated with waters of different quality. Water Resour Res 42:W08401. doi:10.1029/2005WR004802

Grolimund D, Borkovec M, Barmettler K, Sticher H (1996) Colloid-facilitated transport of strongly sorbing contaminants in natural porous media; a laboratory column study. Environ Sci Technol 30:3118–3123

Gumaa GA (1978) Spatial variability of in situ available water. PhD Dissertation, University of Arizona, Tucson, 140 pp

Gwo JP, Jardine PM, Wilson GV, Yeh GT (1995) A multiple-pore-region concept to modeling mass transfer in subsurface media. J Hydrol 164:217–237

Hanselman TA, Graetz DA, Wilkie AC (2003) Manure-borne estrogens as potential environmental contaminants: a review. Environ Sci Technol 37:5471–5478

Hansen S, Jensen HE, Nielsen NE, Svendsen H (1990) DAISY: soil plant atmosphere system model. NPO report no. A 10. The National Agency for Environmental Protection, Copenhagen, 272 pp

Haws NW, Rao PSC, Šimůnek J (2005) Single-porosity and dual-porosity modeling of water flow and transport in subsurface-drained fields using effective field-scale parameters. J Hydrol 313(3–4):257–273

Healy RW (1990) Simulation of solute transport in variably saturated porous media with supplemental information on modifications to the U.S. Geological Survey's computer program VS2DUS. Geological Survey, Water-Resources Investigation Report 90–4025:125

Hendrickx JMH, Dekker LW, Boersma OH (1993) Unstable wetting fronts in water repellent field soils. J Environ Qual 18:109–118

Hendrickx JMH, Flury M (2001) Uniform and preferential flow, mechanisms in the Vadose zone. In: Conceptual models of flow and transport in the fractured Vadose zone. National Research Council, National Academy, Washington, DC, pp 149–187

HydroGeoLogic, Inc (1996) MODFLOW-SURFACT, a comprehensive MODFLOW-based flow and transport simulator, version 2.1, HydroGeoLogic, Inc.

Jacques D, Mouvet JC, Mohanty BP,Vereecken H, Feyen J (1999) Spatial variability of atrazine sorption parameters and other soil properties in a podzoluvisol. J Contamin Hydrol 36: 31–52

Jacques D, Šimůnek J (2005) Multicomponent – variable saturated transport model, description, manual, verification and examples, waste and disposal, SCK•CEN, BLG-998, Mol, Belgium, 79 pp

Jacques D, Šimůnek J, Mallants D, Van Genuchten MTh (2002) Multicomponent transport model for variably-saturated porous media: application to the transport of heavy metals in soils. In: Hassanizadeh SM, Schotting RJ, Gray WG, Pinder GF (eds) Computational methods in water resources, XIVth international conference on computational methods in water resources, June 23–28, 2002, Delft, The Netherlands. Elsevier, Developments in Water Science, vol 47, pp 555–562

Jacques D, Šimůnek J, Mallants D, Van Genuchten MTh (2003) The HYDRUS-PHREEQC multicomponent transport model for variably-saturated porous media: code verification and application, MODFLOW and More 2003: Understanding through Modeling, Conference Proceedings, Poeter E, Zheng Ch, Hill M, Doherty J (eds) International Ground Water Modeling Center, Colorado School of Mines, pp 23–27

Jacques D, Šimůnek J, Mallants D, Van Genuchten MTh (2006) Operator-splitting errors in coupled reactive transport codes for transient variably saturated flow and contaminant transport in layered soil profiles. J Contamin Hydrol 88:197–218

Jacques D, Šimůnek J, Mallants D, Van Genuchten MTh (2008a) Modelling coupled water flow, solute transport and geochemical reactions affecting heavy metal migration in a podzol soil. Geoderma 145:449–491

Jacques D, Šimůnek J, Mallants D, Van Genuchten MTh (2008b) Modeling coupled hydrological and chemical processes in the Vadose zone: a case study of long-term uranium transport following mineral fertilization.Vadose Zone J 7(2):698–711

Jansson P-E, Karlberg L (2001) Coupled heat and mass transfer model for soil-plant-atmosphere systems. Royal institute of technology, Department of Civil and Environmental Engineering, Stockholm, 325 pp

Jarvis NJ (1994) The MACRO model (version 3.1), Technical description and sample simulations. Reports and dissertations 19. Department of Soil Science, Swedish University Agricutural Science, Uppsala, Sweden, 51 pp

Jarvis NJ (1998) Modeling the impact of preferential flow on nonpoint source pollution. In: Selim HM, Ma L (eds) Physical nonequilibrium in soils: modeling and application. Ann Arbor, Chelsea, MI, pp 195–181

Jensen ME, Burman RD, Allen RG (1990) Evapotranspiration and irrigation water requirement. ASCE manuals and reports on enigineering practice, vol 70, ASCE, New York

Jin Y, Jury WA (1996) Characterizing the dependence of gas diffusion coefficient on soil properties. Soil Sci Soc Am J 60:66–71

Jury WA (1982) Simulation of solute transport using a transfer function model. Water Resour Res 18:363–368

Jury WA (1985) Spatial variability of soil physical parameters in solute migration: a critical literature review. EPRI EA-4228, Interim report, EPRI

Jury WA, El Abd H, Collins TM (1983b) Field scale transport of nonadsorbing and adsorbing chemicals applied to the soil surface. NWWA symposium, Las Vegas. National Water Well Assoc, Worthington, OH, pp 203–222

Jury WA, Gardner WA, Gardner WH (1991) Soil physics, 5th edn. Wiley, New York, NY

Jury WA, Horton R (2004) Soil physics, 6th edn. Wiley, New York, NY

Jury WA, Spencer WF, Farmer WJ (1983a) Behavior assessment model for trace organics in soil: I. Description of model. J Environ Qual 12:558–564

Jury WA, Spencer WF, Farmer WJ (1984) Behavior assessment model for trace organics in soil: I. Chemical classification and parameter sensitivity J Environ Qual 13:567–572

Jury WA, Stolzy LH, Shouse P (1982) A field test of the transfer function model for predicting solute transport. Water Resour Res 18:369–375

Kan AT, Tomson MB (1990) Ground water transport of hydrophobic organic compounds in the presence of dissolved organic matter. Environ Toxicol Chem 9:253–263

Köhne S, Lennartz B, Köhne JM, Šimůnek J (2006) Bromide transport at a tile-drained field site: experiment, one- and two-dimensional equilibrium and non-equilibrium numerical modeling. J Hydrol 321(1–4):390–408

Köhne JM, Köhne S, Šimůnek J (2009a) A review of model applications for structured soils: a) water flow and tracer transport. J Contamin Hydrol, Special issue “flow domains” 104(1–4): 4–35. doi:10.1016/j.jconhyd.2008.10.002

Köhne JM, Köhne S, Šimůnek J (2009b) A review of model applications for structured soils: b) pesticide transport. J Contamin Hydrol, Special issue "flow domains" 104(1–4):36–60. doi:10.1016/j.jconhyd.2008.10.003

Köhne JM, Mohanty B, Šimůnek J, Gerke HH (2004) Numerical evaluation of a second-order water transfer term for variably saturated dual-permeability models. Water Resour Res 40. doi:10.1029/2004WR00385

Kosugi K (1996) Lognormal distribution model for unsaturated soil hydraulic properties. Water Resour Res 32(9):2697–2703

Kung K-JS (1990) Preferential flow in a sandy vadose zone. 2. Mechanism and implications. Geoderma 46:59–71

Langmuir I (1918) The adsorption of gases on plane surfaces of glass, mica, and platinum. J Am Chem Soc 40:1361–1403

Leij FJ, Bradford SA (1994) 3DADE: a computer program for evaluating three-dimensional equilibrium solute transport in porous media. Research Report No. 134 US Salinity Laboratory, USDA, ARS, Riverside, CA

Leij FJ, Toride N (1997) N3DADE: a computer program for evaluating nonequilibrium three-dimensional equilibrium solute transport in porous media. Research Report No. 143 US Salinity Laboratory, USDA, ARS, Riverside, CA

Leij FJ, Priesack E, Schaap MG (2000) Solute transport modeled with Green's functions with application to persistent solute sources. J Contamin Hydrol 41:155–173

Leij FJ, Schaap MG, Arya LM (2002) Indirect methods. In: Dane JH, Topp GC (eds) Methods of soil analysis, part 1, physical methods, chapter 3.6.3, 3rd edn. SSSA, Madison, WI, pp 1009–1045

Lenhard RJ, Oostrom M, Dane JH (2002) Multi-fluid flow. In: Dane JH, Topp GC (eds) Methods of soil analysis, part 4, physical methods, SSSA book series 5, Soil Sci Soc Am, Madison, WI, pp 1537–1619

Lichtner PC (1996) Continuum formulation of multicomponent-multiphase reactive transport. In: Lichtner PC, Steefel CI, Oelkers EH (eds) Reactive transport in porous media, Reviews in mineralogy. Chapter 1. Mineralogical Society of America, vol 34, pp 1–81

Lichtner PC (2000) FLOWTRAN user's manual. Los Alamos National Laboratory Document. NM

Lichtner PC, Seth MS (1996) User's manual for MULTIFLO: Part II MULTIFLOW 1.0 and GEM 1.0. Multicomponent-multiphase reactive transport model. CNWRA 96–010, Center for Nuclear Waste Regulatory Analyses, San Antonio, TX

Lindqvist R, Enfield CG (1992) Biosorption of dichlorodiphenyltri-chloroethane and hexachlorobenzene in groundwater and its implications for facilitated transport. Appl Environ Microbiol 58:1811–1818

Liu CW, Narasimhan TN (1989) Redox-controlled multiple-species reactive chemical transport. 1. Model development. Water Resour Res 25(5):869–882

Liu C, Szecsody JE, Zachara JM, Ball WP (2000) Use of the generalized integral transform method for solving equations of solute transport in porous media. Adv Water Resour 23(5):483–492

Mallants D, Jacques D, Vanclooster M, Diels J, Feyen J (1996b) A stochastic approach to simulate water flow in a macroporous soil. Geoderma 70:299–324

Mallants D, Mohanty BP, Diederik J, Feyen J (1996c) Spatial variability of hydraulic properties in a multi-layered soil profile. Soil Sci 161(3):167–181

Mallants D, Tseng P-H, Toride NA (1997) Timmerman, and Feyen J. Evaluation of multimodal hydraulic functions in characterizing a heterogeneous field soil. J Hydrol 195:172–199

Mallants D, Vanclooster M, Feyen J (1996a) A transect study on solute transport in a macroporous soil. Hydrol Process 10:55–70

Mallants D, Volckaert G, Marivoet J (1999) Sensitivity of protective barrier performance to changes in rainfall rate. Waste Manage 19:467–475

Magee BR, Lion LW, Lemley AT (1991) Transport of dissolved organic macromolecules and their effect on the transport of phenanthrene in porous media. Environ Sci Technol 25:323–331

Mangold DC, Tsang CF (1991) A summary of subsurface hydrological and hydrochemical models. Rev Geophys 29(1):51–79

Mansfeldt T, Leyer H, Barmettler K, Kretzschmar R (2004) Cyanide leaching from soil developed from coking plant purifier waste as influenced by citrate. Vadose Zone J 3:471–479

Massoudieh A, Ginn TR (2009) Modeling colloid-enhanced contaminant transport in stormwater infiltration basin best management practices. Vadose Zone J 7:1261–1268

Mayer U, Frind EO, Blowes DW (2002) Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. Water Resour Res 38:1174

MDH Engineered Solutions Corp (2003) Evaluation of computer models for predicting the fate and transport of salt in soil and groundwater, Phase I report. Pub. No: T/403, ISBN No. 0-7785-2493-0 (printed edition), ISBN No. 0-7785-2494-9 (on-line edition)

Millington RJ, Quirk JM (1961) Permeability of porous solids. Trans Faraday Soc 57:1200–1207

Moldrup O, Poulsen TG, Schjonning P, Olesen T, Yamaguchi T (1998) Gas permeability in undisturbed soils: measurements and predictive models. Soil Sci 163:180–189

Monteith JL (1965) Evaporation and environment. In: Fogg GE (ed) Symposium of the society for experimental biology. The state and movement of water in living organisms, vol 19. Academic, NY, pp 205–234

Mualem Y (1976) A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour Res 12(3):513–522

Narasimhan TN (2005) Buckingham, 1907; An appreciation. Vadose Zone J 4:434–441

Nielsen DR, Biggar JW, Erh KT (1973) Spatial variability of field measured soil water properties. Hilgardia 42:215–259

Nielsen DR, Van Genuchten MTh, Biggar JW (1986) Water flow and solute transport processes in the unsaturated zone. Water Resour Res 18(9):89S-108S

Noell AL, Thompson JL, Corapcioglu MY, Triay IR (1998) The role of silica colloids on facilitated cesium transport through glass bead columns and modeling. J Contamin Hydrol 31: 23–56

Pachepsky YA, Rawls WJ (eds) (2004) Development of pedotransfer functions in soil hydrology. Elsevier, Amsterdam

Pachepsky YA, Rawls WJ, Lin HS (2006) Hydropedology and pedotransfer functions. Geoderma 131:308–316

Parker JC, Van Genuchten MTh (1984) Determining transport parameters from laboratory and field tracer experiments. Bull 84–83, Va Agric Exp St, Blacksburg, VA

Parkhurst DL, Appelo CAJ (1999) User's guide to PHREEQC (version 2) – a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. Water resources investigation, report 99-4259, Denver, CO, USA, 312 pp

Penman HL (1948) Natural evaporation from open water, bare soil, and grass. Proc Roy Soc Lon Z193:120–156

Pot V, Šimůnek J, Benoit P, Coquet Y, Yra A, Martínez-Cordón M-J (2005) Impact of rainfall intensity on the transport of two herbicides in undisturbed grassed filter strip soil cores. J. Contamin Hydrol 8:63–88

Pruess K (1991) Tough2 – a general-purpose numerical simulator for multiphase fluid and heat flow. Rep LBL-29400, Lawrence Berkeley Laboratory, Berkeley, CA

Pruess K (2004) The TOUGH codes -A family of simulation tools for multiphase flow and transport processes in permeable media. Vadose Zone J 3:738–746

Pruess K, Wang JSY (1987) Numerical modeling of isothermal and non-isothermal flow in unsaturated fractured rock – a review, In: Evans DD, Nicholson TJ (eds) Flow and transport through unsaturated fractured rock. Geophysics monograph, vol 42. American Geophysical Union, Washington, DC, pp 11–18

Rathfelder KM, Lang JR, Abriola LM (2000) A numerical model (MISER) for the simulation of coupled physical, chemical and biological processes in soil vapor extraction and bioventing systems. J Contamin Hydrol 43(3–4):239–270

Richards LA (1931) Capillary conduction of fluid through porous mediums. Physics 1:318–333

Ritsema CJ, Dekker LW (2000) Preferential flow in water repellent sandy soils: principles and modeling implications. J Hydrol 231–232(29):308–319

Russo D, Bresler E (1981) Soil hydraulic properties as stochastic processes: 1. Analysis of field spatial variability. Soil Sci Soc Am J 45:682–687

Scanlon BR, Christman M, Reedy R, Porro I, Šimůnek J, Flerchinger G (2002) Intercode comparisons for simulating water balance of surficial sediments in semiarid regions. Water Resour Res 38(12):1323. doi:10.1029/2001WR001233, 59.1–59.16

Schaap MG, Leij FJ (2000) Improved prediction of unsaturated hydraulic conductivity with the Mualem-Van Genuchten model. Soil Sci Soc Am J 64:843–851

Schaap MG, Leij FJ, Van Genuchten MTh (2001) Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J Hydrol 251:163–176

Schaerlaekens J, Mallants D, Šimůnek J, Van Genuchten MTh, Feyen J (1999) Numerical simulation of transport and sequential biodegradation of chlorinated aliphatic hydrocarbons using CHAIN_2D. J Hydrol Process 13(17):2847–2859

Schijven J, Šimůnek J (2002) Kinetic modeling of virus transport at field scale. J Contamin Hydrol 55(1-2):113-135

Selim HM, Büchter B, Hinz C, Ma L (1992) Modeling the transport and retention of cadmium in soils: multireaction and multicomponent approaches. Soil Sci Soc Am J 56:1004–1015

Selim HM, Davidson JM, Mansell RS (1976) Evaluation of a two-site adsorption-desorption model for describing solute transport in soils. Paper presented at proceedings. Summer Computer Simulation Conference, Nat Sci Found, Washington, DC

Seuntjens P (2000) Reactive solute transport in heterogeneous porous medium: cadmium leaching in acid sandy soils. Ph.D., University of Antwerp, Belgium

Seuntjens P, Tirez K, Šimůnek J, Van Genuchten MTh, Cornelis C, Geuzens P (2001) Aging effects on cadmium transport in undisturbed contaminated sandy soil columns. J Environ Qual 30:1040–1050

Šimůnek J (2005) Models of water flow and solute transport in the unsaturated zone. In: Anderson MG (ed) The encyclopedia of hydrological sciences. Wiley, Chichester, England, pp 1171–1180

Šimůnek J, Changming H, Pang JL, Bradford SA (2006a) Colloid-facilitated transport in variably-saturated porous media: numerical model and experimental verification. Vadose Zone J 5:1035–1047

Šimůnek J, Jarvis NJ, Van Genuchten MTh, Gärdenäs A (2003) Nonequilibrium and preferential flow and transport in the vadose zone: review and case study. J Hydrol 272:14–35

Šimůnek J, Jacques D, Van Genuchten MTh, Mallants D (2006b) Multicomponent geochemical transport modeling using the HYDRUS computer software packages. J Am Water Resour Assoc 42(6):1537–1547

Šimůnek J, Suarez DL (1993) Modeling of carbon dioxide transport and production in soil: 1. Model development. Water Resour Res 29(2):487–497

Šimůnek J, Suarez DL (1994) Two-dimensional transport model for variably saturated porous media with major ion chemistry. Water Resour Res 30(4):1115–1133

Šimůnek J, Suarez DL (1997) Sodic soil reclamation using multicomponent transport modeling. ASCE J Irrig Drain Eng 123(5):367–376

Šimůnek J, Šejna M, Van Genuchten MTh (1999b) The HYDRUS-2D software package for simulating two-dimensional movement of water, heat, and multiple solutes in variably saturated media. Version 2.0, IGWMC – TPS – 53, International Ground Water Modeling Center, Colorado School of Mines, Golden, CO, 251 pp

Šimůnek J, Šejna M, Van Genuchten MTh (1998a) The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple in variably-saturated media. Version 2.0, IGWMC – TPS – 70, International Ground Water Modeling Center, Colorado School of Mines, Golden, CO, 202 pp

Šimůnek J, Suarez DL, Šejna M (1996) The UNSATCHEM software package for simulating one-dimensional variably saturated water flow, heat transport, carbon dioxide production and transport, and multicomponent solute transport with major ion equilibrium and kinetic chemistry. Version 2.0. Research Report No. 141 US Salinity Laboratory, USDA, ARS, Riverside, CA, 186 pp

Šimůnek J, Valocchi AJ (2002) Geochemical transport. In: Dane JH, Topp GC (eds) Methods of soil analysis, part 1, physical methods, chapter 6.9, 3rd edn. SSSA, Madison, WI, pp 1511–1536

Šimůnek J, Van Genuchten MTh (2006) Contaminant transport in the unsaturated zone: theory and modeling, chapter 22. In: Delleur J (ed) The handbook of groundwater engineering, 2nd edn. CRC, Boca Raton, FL, pp 22.1–22.46

Šimůnek J, Van Genuchten MTh (2008) Modeling nonequilibrium flow and transport with HYDRUS. Vadose Zone J, Special issue "Vadose zone modeling" 7:782–797

Šimůnek J, Van Genuchten MTh, Šejna M (2005) The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 3.0, HYDRUS Software Series 1, Department of Environmental Sciences, University of California Riverside, Riverside, CA, 270 pp

Šimůnek J, Van Genuchten MTh, Šejna M (2006c) The HYDRUS software package for simulating two- and three-dimensional movement of water, heat, and multiple solutes in variably-saturated media, Technical manual, version 1.0, PC Progress, Prague, Czech Republic, 241 pp

Šimůnek J, Van Genuchten MTh, Sejna M (2008) Development and applications of the HYDRUS and STANMOD software packages and related codes. Vadose Zone J, Special issue "Vadose zone modelling" 7(2):587–600

Šimůnek J, Van Genuchten MTh, Šejna M, Toride N, Leij FJ (1999a) The STANMOD computer software for evaluating solute transport in porous media using analytical solutions of convection-dispersion equation. Versions 1.0 and 2.0, IGWMC – TPS – 71, International Ground Water Modeling Center, Colorado School of Mines, Golden, CO, 32 pp

Šimůnek J, Wendroth O, Van Genuchten MTh (1998b) A parameter estimation analysis of the evaporation method for determining soil hydraulic properties. Soil Sci Soc Am J 62(4): 894–905

Šimůnek J, Wendroth O, Wypler N, Van Genuchten MTh (2001) Nonequilibrium water flow characterized from an upward infiltration experiment. Eur J Soil Sci 52(1):13–24

Smolders E, McLaughlin M (1996) Effect of Cl on Cd uptake by Swiss Chard in nutrient solutions. Plant Soil 179:57–64

Sposito G, Barry DA (1987) On the Dagan model of solute transport I groundwater: foundational aspects. Water Resour Res 23:1867–1875

Steefel CI, MacQuarrie KTB (1996) Approaches to modeling of reactive transport in porous media. In: Lichtner PC, Steefel CI, Oelkers EH (eds) Reactive transport in porous media. Reviews in mineralogy. Chapter 2, vol 34, Mineralogical Society of America, pp 83–129

Steefel CI, Yabusaki SB (1996) OS3D/GIMRT, software for multicomponent-multidimensional reactive transport. User manual and programmer's guide, PNL-11166, Pacific Northwest National Laboratory, Richland, WA

Thibault DH, Sheppard MI, Smith PA (1990), A critical compilation and review of default soil solid/liquid partition coefficients, K_d, for use in environmental assessments. AECL-10125

Thiele-Bruhn S (2003) Pharmaceutical antibiotic compounds in soils – a review. J Plant Nutr Soil Sci 166:145–167

Tolls J (2001) Sorption of veterinary pharmaceuticals in soils: a review. Environ Sci Technol 35:3397–3406

Toride N, Leij FJ, Van Genuchten MTh (1993) A comprehensive set of analytical solutions for nonequilibrium transport with first-order decay and zero-order production. Water Resour Res 29(7):2167–2182

Toride N, Leij FJ, Van Genuchten MTh (1995) The CXTFIT code for estimating transport parameters from laboratory or field tracer experiments. Version 2.0, Research report no. 137 US Salinity Laboratory, USDA, ARS, Riverside, CA

Totsche KU, Kögel-Knaber I (2004) Mobile organic sorbent affected contaminant transport in soil: numerical case studies for enhanced and reduced mobility. Vadose Zone J 3:352–367

Ungs M, Wilkinson K, Summers K (1998) Users guide for MYGRT, Version 3.0: software for simulating migration of organic and inorganic chemicals in grioundwater. EPRI Report TR-111748, Electric Power Research Institute, Palo Alto, CA

van Dam JC, Huygen J, Wesseling JG, Feddes RA, Kabat P, van Valsum PEV, Groenendijk P, van Diepen CA (1997) Theory of SWAP, version 2.0. Simulation of water flow, solute transport and plant growth in the soil- water- atmosphere- plant environment. Department Water Resources, WAU, Report 71, DLO Winand Staring Centre, Wageningen, Technical Document 45

Vanderborght J, Kasteel R, Herbst M, Javaux M, Thiery D, Vanclooster M, Mouvet C, Vereecken H (2005) A set of analytical benchmarks to test numerical models of flow and transport in soils. Vadose Zone J 4:206–181

Van Genuchten MTh (1980a) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am J 44:892-898

Van Genuchten MTh (1980b) Determining transport parameters from solute displacement experiments. Research report no. 118 US Salinity Laboratory, USDA, ARS, Riverside, CA

Van Genuchten MTh (1981) Non-equilibrium transport parameters from miscible displacement experiments. Research report no. 119 US Salinity Laboratory, USDA, ARS, Riverside, CA

Van Genuchten MTh (1987) A numerical model for water and solute movement in and below the root zone. Research Report no. 121 US Salinity Laboratory, USDA, ARS, Riverside, CA

Van Genuchten MTh, Parker JC (1984) Boundary conditions for displacement experiments through short laboratory soil columns. Soil Sci Soc Am J 48(4):703-708

Van Genuchten MTh, Šimůnek J (1996) Evaluation of pollutant transport in the unsaturated zone. In: Rijtema PE, Eliáš V (eds),Regional approaches to water pollution in the environment, NATO ASI series: 2. Environment. Kluwer, Dordrecht, The Netherlands, pp 139–172

Van Genuchten MTh, Wagenet RJ (1989) Two-site/two-region models for pesticide transport and degradation: theoretical development and analytical solutions. Soil Sci Soc Am J 53: 1303–1310

Vinten AJA, Yaron B, Nye PH (1983) Vertical transport of pesticides into soil when adsorbed on suspended particles. J Agric Food Chem 31:662–664

Voegelin A (2001) Competetive sorption and transport of heavy metal cations in soils. Diss ETH No. 14101, Zürich, Switzerland

Voegelin A, Vulava V, Kretzschmar R (2001) Reaction-based model describing competitive sorption and transport of Cd, Zn, and Ni in an acidic soil. Environ Sci Technol 35: 1651–1657

Von Gunten HR, Weber VE, Krahenbuhl U (1998) The reactor accident at Chernobyl: a possibility to test colloid-controlled transport of radionuclides in a shallow aquifer. J Contamin Hydrol 2:237–247

Wagenet RJ, Hutson JL (1987) LEACHM: Leaching Estimation And CHemistry Model, A process-based model of water and movement, transformations, plant uptake and chemical reactions in the unsaturated zone. Continuum 2. Water Resour Inst, Cornell University, Ithaca, New York

Wang HF, Anderson MP (1982) Introduction to groundwater modeling: finite difference and finite element methods. Academic, San Diego

White N, Zelazny LW (1986) Charge properties in soil colloids. In: Sparks DL (ed) Soil physical chemistry. CRC, Boca Raten, FL, pp 39–81

Wheatcraft SW, Cushman JH (1991) Hierarchical approaches to transport in heterogeneous porous media. Rev Geophys 29(Suppl):261–267

Wild A, Babiker IA (1976) The asymmetric leaching pattern of NO_3 and Cl in a loamy sand under field conditions. J Soil Sci 27:460–466

Willardson LS, Hurst RL (1965) Sample size estimates in permeability studies. J Irrigation Drain Div, ASCE, 91:1–9

Wilson GV, Albright WH, Gee GW, Fayer MJ, Ogan BD (1999) Alternative cover assessment project, Phase 1 Report. Final technical report prepared by Desert research institute, NV and Pacific Northwest National Laboratory, WA for U.S. Environmental Protection Agency under Contract No. 44-0000-2038, 202 pp

Yeh GT (1981) Analytical transient one-, two-, and three-dimensional simulation of waste transport in the aquifer system, ORNL-5602, Oak Ridge National Laboratory, Oak Ridge, TN

Yeh GT, Tripathi VS (1990) HYDROGEOCHEM: a coupled model of hydrological and geochemical equilibrium of multicomponent systems. Rep ORNL-6371, Oak Ridge Natl Lab, Oak Ridge, TN

Yeh GT, Salvage KM, Gwo JP, Zachara JM, Szecsody JE (1998) HYDROBIOGEOCHEM: a coupled model of hydrological transport and mixed biochemical kinetic/equilibrium reactions in saturated-unsaturated media. Rep. ORNL/TM-13668, Oak Ridge Natl Lab, Oak Ridge, TN

Zyvoloski GA, Robinson BA, Dash ZV, Trease LL (1997) Summary of the models and methods for the FEHM application – a finite element heat- and mass-transfer code. Los Alamos National Laboratory Report LA-13307-MS, NM

Chapter 19
Contaminant Fate and Reactive Transport in Groundwater

Massimo Rolle, Ulrich Maier, and Peter Grathwohl

Abstract Understanding the complex, interacting processes that determine the fate and transport of contaminants in groundwater is a major challenge for evaluating and predicting risks to clean water, human and ecological receptors and for designing effective remediation plans. Different physical and biogeochemical processes including advection, hydrodynamic dispersion, dissolution, sorption and biodegradation affect the migration of contaminants in saturated porous media like a groundwater system. In this chapter an overview of these processes is presented together with the basic theory on contaminant transport modeling, which represents an essential tool for a quantitative description of contaminant migration in the subsurface. Numerical simulations of typical contamination scenarios are presented, with the main goal of identifying the influence of different parameters on contaminant fate and transport such as transverse dispersivity, thickness and strength of the contamination source, recharge, biodegradation rates and mixing enhancement through flow focusing in high permeability zones. These numerical simulations are complemented by two examples, i.e. the reactive transport of toluene from a LNAPL source and a field study, where ammonium is continuously released from a leaking landfill to the underlying aquifer. The principal processes at the landfill site have been quantitatively integrated into the framework of a two-dimensional reactive transport model.

Contents

M. Rolle (✉)
University of Tübingen, Tübingen, Germany
e-mail: massimo.rolle@uni-tuebingen.de

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_19,

19.1 Introduction

Contamination of groundwater resources is one of the most urgent environmental issues throughout the world. Groundwater is a predominant source for drinking water, and it is also extensively used for agricultural food production (through irrigation) and many industrial purposes. Thus, the deterioration of its quality is detrimental for human health and for the ecological systems.

Principal sources of groundwater contamination are various industrial and domestic activities such as leaking underground storage tanks, dry-cleaning, use of pesticides and herbicides in agriculture, exploitation of oil and gas fields, traffic, landfills, et cetera. A wide variety of contaminants is released into groundwater systems including mono- and polycyclic aromatic hydrocarbons (BTEX, PAHs), chlorinated solvents, polychlorinated biphenyls (PCBs), phenols, heavy metals, ammonia, nitrate, pesticides, pharmaceuticals and radionuclides (Fetter 1993). The understanding of the fate and transport of contaminants in groundwater systems has become a significant issue in environmental science during the last decades and is still a dynamic and active field of applied research. Complex physico-bio-geochemical processes, such as advection and dispersion, sorption, ion exchange, volatilization, chemical reactions and biodegradation are involved in the fate of contaminants in the subsurface. The characterization and quantification of these processes play a crucial role for the understanding of the contaminants' environmental fate, for risk-based groundwater quality assessment and for the application of adequate groundwater remediation technologies. In particular, in the last years innovative remediation methods (e.g. bioremediation, monitored Natural Attenuation, subsurface reactive barriers, et cetera) have emerged and proved to be more effective than traditional highly intrusive technologies, such as pump and treat, on the long term. These new remediation methods require a deeper knowledge of the complex and interacting subsurface processes in order to predict the fate of contaminants in the subsurface.

In the following an overview of the principal processes determining the fate and transport of contaminants in groundwater is presented together with an introduction to the basic theory on reactive transport models. Besides, some applications including a simplified reactive transport scenario modeling and a field study are illustrated. These applications are focused on the investigation of the fate and transport of two kinds of contaminants frequently found at contaminated sites: monoaromatic petroleum hydrocarbons (BTEX) and ammonium (NH_4^+).

19.2 Basic Theory on Contaminant Transport

The first step when dealing with contaminated groundwater is the definition of a conceptual model on flow and contaminant transport at a specific site. Many pieces of information need to be integrated in the conceptual model including the knowledge of the physical configuration of the aquifer (i.e. geometry, areal extent, depth of groundwater table, thickness and geological characteristics of the different formations, hydraulic boundary conditions), hydraulic properties (i.e.flow direction, hydraulic gradient, permeabilities, storage coefficients, recharge rates), transport properties (i.e. dispersion coefficients, effective porosity), geochemistry of the background groundwater and interaction with the solid matrix, physico-chemical properties of the contaminants (i.e. density, solubility, partition coefficients, vapor pressure), information about the contaminant release scenario (i.e. suspected source location, mass and duration of the leakage) and about potential targets (i.e. human and/or ecological receptors). The development of the conceptual model should lead to a hypothesis on the interaction of sources, migration pathways and receptors. Figure 19.1 shows a sketch of a conceptual representation of contaminant release and migration in a shallow unconfined aquifer underlying an industrial site.

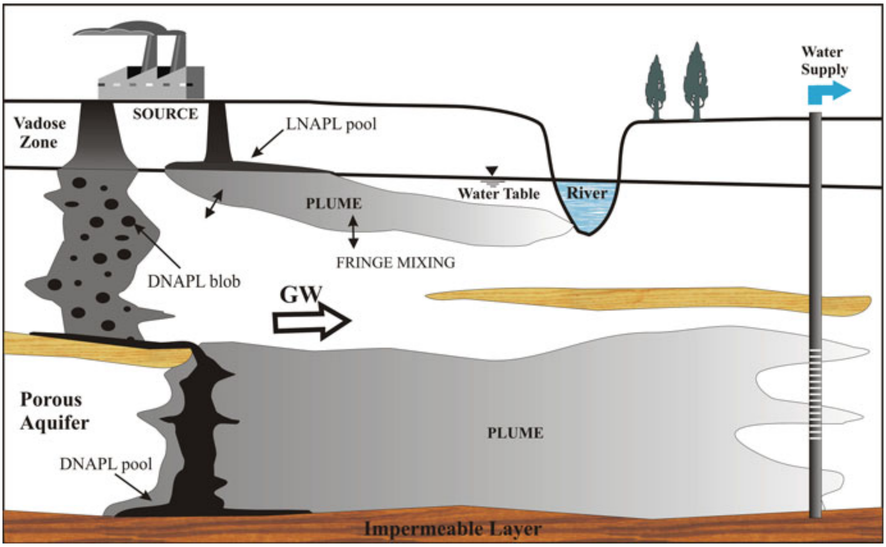


Fig. 19.1 Schematic representation of groundwater contamination and impact on sensitive receptors

The main migration pathway of contaminants in aquifer systems implies their transport as dissolved species forming a plume in the groundwater flow direction. The contaminant plumes can impact some sensitive receptors such as surface water bodies or water supply wells. Different physical and biogeochemical processes affect the migration of contaminant plumes including the dissolution from NAPL (Non-Aqueous Phase Liquid) sources, advective transport, longitudinal and transverse dispersion, sorption to the solid matrix of the aquifer and biological degradation.

19.2.1 Contamination Sources and Plume Formation

At the majority of contaminated groundwater sites, organic contaminants, such as fuels and chlorinated solvents, are present in the form of NAPLs that have migrated into the subsurface. Dissolution of NAPLs, the process by which contaminants transfer from the NAPL into the aqueous phase, is one of the most important processes that determines the formation of dissolved contaminant plumes. NAPLs are organic, non-miscible liquids in the subsurface. The dynamic of NAPLs migration in the subsurface can be very complex (Pankow and Cherry 1996); a schematic representation is illustrated in Fig. 19.1. When introduced into the vadose zone, they initially stay in a connected body of continuous NAPL (also called free-phase NAPL, or mobile NAPL, or non-residual NAPL or free product) that migrates through the subsurface, invading pore spaces, fractures and any preferential pathways in the subsurface by displacing water or air in the unsaturated zone. At the water table a major difference in behavior is observed between organic liquids less dense than water (LNAPLs) and those denser than water (DNAPLs). LNAPLs will float on the top of the water table, forming a thin, pancake-like layer that spreads across the water surface. Being denser than water, DNAPLs will penetrate the water table, displace the water from the pores of the aquifer and migrate into the saturated zone and finally will spread laterally where changes in the matrix permeability (low permeability barriers) inhibit vertical NAPL migration. Once the contaminant supply ceases, the NAPL begins to spread laterally and discrete, isolated blobs start to form (residual NAPL).

Although NAPLs can be present as a single contaminant (e.g. trichloroethene), they usually are composed of a mixture of contaminants. Typical LNAPLs such as gasoline and diesel (density in the range 0.8–0.9 kg L^{-1}) are composed of hundreds of different contaminants with a wide variety of chemical and physical properties. Many DNAPLs like industrial solvents, coal tar, creosote, et cetera (density range 1.01–1.2 kg L^{-1}) are complex mixtures of different contaminants as well.

NAPL dissolution in the groundwater depends on different factors such as the interfacial area between the organic phase and water, the extent and morphology of the source, the groundwater flow velocity, the solubility of individual contaminants and the composition of the NAPL source. Eberhardt and Grathwohl (2002) showed that for complex, multicomponent coal tar sources the equilibrium or saturation concentration ($C_{i,sat}$) of a contaminant in water (e.g. at the NAPL/water-interface) can be described by the Raoult's law:

$$C_{i,sat} = \chi_{i,o}\gamma_{i,o}S_i \tag{19.1}$$

where $\chi_{i,o}$ [dimensionless], $\gamma_{i,o}$[dimensionless] and S_i [M L^{-3}], denote the molar fraction of contaminant *i* in the organic contaminant mixture, the activity coefficient of *i* and the aqueous solubility of the individual contaminant *i* (pure substance) [M L^{-3}], respectively.

For mass transfer into groundwater a film diffusion model can be used to describe the dissolution rates F_b [M L^{-2}T^{-1}] from NAPLs trapped as blobs or ganglia in a porous medium (Grathwohl 1998):

$$F_b = \frac{D_{aq}}{\delta}(C_0 - C) \tag{19.2}$$

where D_{aq} is the aqueous diffusion coefficient [L^{2}T^{-1}], δ is the film thickness [L], C is the contaminant concentration in the aqueous phase and C_0 is the contaminant concentration at the interface. The ratio D_{aq}/δ is often indicated as mass transfer coefficient k [L T^{-1}]. This parameter is generally unknown and empirical correlations are often used to estimate it based on dimensionless constants such as the Sherwood, Schmidt and Reynolds numbers (e.g. Imhoff et al. 1993; Miller et al. 1990). The overall dissolution rate also depends on the interfacial area of the NAPL. Generally the length of mass transfer zones in groundwater flowing through areas with residual NAPL existing as blobs or ganglia is rather short in the range of centimeters and decimeters indicating that equilibrium contaminant concentrations in this scenario are reached very rapidly.

For NAPL entrapped as connected free phase on top of low permeability layers and impervious formations ("DNAPL pools") or floating on the water table ("LNAPL pools") the dissolution rates depend on the contact time between the aqueous and coherent organic phases, the pool dimensions and the transverse vertical dispersion. In a first approximation the dissolutions rate can be described as (see Grathwohl 1998 for further detail):

$$F_p = 2\,C_0 n\sqrt{\frac{D}{\pi \cdot t_c}}L_p B_p \tag{19.3}$$

where C_0 is the equilibrium concentration at the interface between organic and water phase, n is the porosity through which the water can flow, D [L^{2}T^{-1}] is the vertical transverse dispersion coefficient, t_c [T] is the contact time and L_p and B_p [L] are the pool length and width, respectively.

Because of the less favorable surface to volume ratio, contaminant dissolution rates from NAPL pools are significantly lower that the ones from disconnected NAPL blobs resulting in longer time and higher resistance to remediation.

19.2.2 Advection

Advection refers to the transport of contaminants by the bulk movement of groundwater and it often represents the most important process driving the down-gradient migration of dissolved contaminants in aquifer systems. The motion of groundwater can be described by the experimentally derived Darcy's law, which, in three-dimensions, can be expressed as:

$$q_i = -K_{ij}\frac{\partial h}{\partial x_i} \tag{19.4}$$

where q_i [L T^{-1}] is the specific discharge or Darcy velocity (discharge per unit cross sectional bulk area), K_{ij} [L T^{-1}] is the hydraulic conductivity tensor, x_i is the spatial coordinate and h [L] is the hydraulic head. The average linear velocity of the flow can be defined as the flux of water across the unit cross sectional area of pore space and is described as the ratio between the specific discharge and the porosity:

$$v_i = \frac{q_i}{n} \tag{19.5}$$

The average linear velocity should not be confused with the local (microscopic) velocity of the fluid or with the average velocity at which water molecules travel along a flow path, which is greater than the average linear velocity because of the tortuosity of the porous medium.

Contaminant transport by advection alone is often described as plug flow and yields a sharp contaminant concentration front advancing along the groundwater flow direction. This concept is illustrated in Fig. 19.2 where the one-dimensional advective transport from an instantaneous and a continuous contaminant release is depicted. After a time t_1 the contaminant front reaches the distance x_1, whereas at a later time t_2 the front has moved to a farther distance x_2.

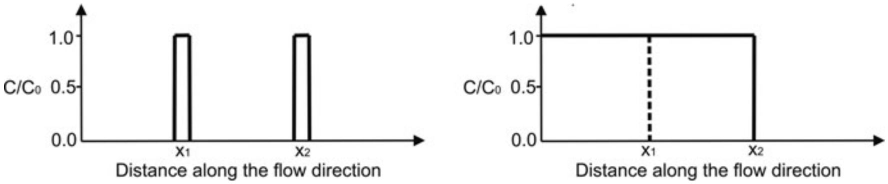


Fig. 19.2 Relative concentration as a function of distance in the flow direction, illustrating the advective movement of a contaminant from an instantaneous (*left*) and a continuous release (*right*). The position x_1 is reached at time t_1 and the position x_2 is reached at time t_2

19.2.3 Hydrodynamic Dispersion

Hydrodynamic dispersion is the process whereby a contaminant plume spreads out from the main direction of groundwater flow and results in dilution with an increasing volume of the flow domain, resulting in a contaminated zone broader than expected from the effect of advective flow alone. Two processes contribute to hydrodynamic dispersion: molecular diffusion and mechanical dispersion.

Aqueous molecular diffusion describes the mass transport of a contaminant due to the random thermal motion of molecules and atoms, also known as Brownian motion. The mass flux of contaminant F per unit cross-sectional area [M L^{-2} T^{-1}] depends on the concentration gradient and can be expressed by Fick's first law which, for a one dimensional case, is written as:

$$F = -D_{aq}\frac{\partial C}{\partial x} \tag{19.6}$$

where D_{aq} is the aqueous diffusion coefficient [$\mathrm{L}^2\mathrm{T}^{-1}$].

In porous media aqueous diffusion is hindered by the tortuous nature of the pores, the diminished cross sectional area available for diffusion and by the size of the pores. Therefore, an effective pore diffusion coefficient can be defined as:

$$D_e = \frac{D_{aq} n \delta}{\tau_f} \tag{19.7}$$

where δ is a dimensionless factor (≤ 1) accounting for pore constrictivity and τ_f is the tortuosity defined as the square ratio of the effective path length in the pore l_e [L] and the shortest distance l [L]:

$$\tau_f = \left(\frac{l_e}{l}\right)^2 > 1 \tag{19.8}$$

In most practical cases the pore size distribution and tortuosities are unknown and only the porosity is known. Therefore, the effective diffusion coefficient is often described, by empirical correlations, as a function of the aqueous diffusion coefficient and the porosity (Grathwohl 1998):

$$D_e = D_{aq} n^m \tag{19.9}$$

where m is an empirical exponent (mostly close to 2). Under transient transport conditions D_e is divided by a capacity factor which accounts for the storage of the contaminant in the porous media (i.e. the porosity for tracers). Molecular diffusion, acting in the direction of a decreasing concentration gradient, tends to equalize the concentration differences along discrete flow paths and is essential for mixing of solutes across different flow paths.

Mechanical dispersion can be defined as the contaminant spreading caused by local variation of the flow velocity. At the microscopic scale, these velocity variations depend on the tortuosity, on the size of the pores and on the variable friction within an individual pore, with faster flow close to pore axis and slower flow in proximity of the solid particle surface (Fig. 19.3).

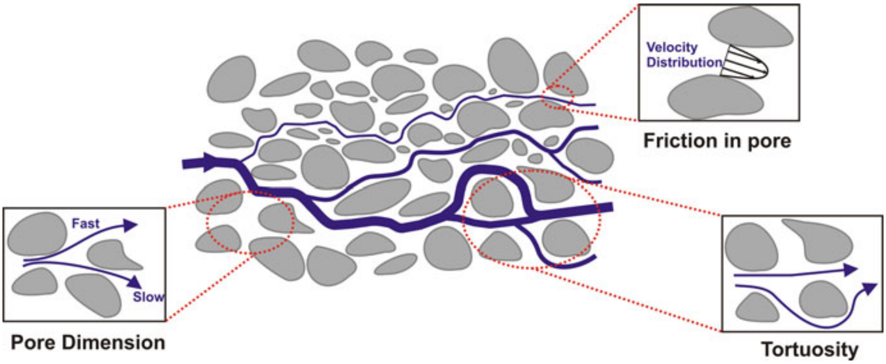


Fig. 19.3 Dispersion in a porous medium, illustrating the influence of pore size, friction in the pores and tortuosity

At the field scale also macroscopic heterogeneities, particularly of aquifer properties such as hydraulic conductivity and porosity, have to be considered in addition to pore-scale heterogeneity. Macroscopic heterogeneities cause the flow to be focused into zones of higher conductivity, to diverge around less permeable formations and to be refracted in the presence of conductivity boundaries. Therefore, the moving contaminant will experience variations from the average linear velocity also at the macroscale.

Mechanical dispersion can occur both, along the flow direction (longitudinal dispersion) and normal to the flow path (transverse dispersion). The dispersive mass flux is often described by a Fickian type law; the hydromechanical dispersion coefficient is often described as the product of the dispersivity and the average linear velocity. The overall process of hydrodynamic dispersion, taking into account the simultaneous occurrence of molecular diffusion and hydromechanical dispersion, can be described by means of hydrodynamic dispersion coefficients, as follows:

$$D_L = D_p + \alpha_L v_i \tag{19.10}$$

$$D_T = D_p + \alpha_T v_i \tag{19.11}$$

where D_L [L^2 T^{-1}] and α_L [L] are the longitudinal dispersion coefficient and dispersivity, D_T [L^2 T^{-1}] and α_T [L] are the transverse dispersion coefficient and dispersivity and D_p is the pore diffusion coefficient (e.g. D_e/n). Recent research on transverse dispersion points towards a nonlinear relationship between the hydromechanical dispersion coefficient and the flow velocity (Chiogna et al. 2010; Klenk and Grathwohl 2002; Olsson and Grathwohl 2007).

The effect of hydrodynamic dispersion on the movement of a contaminant front in one dimension is depicted in Fig. 19.4. It can be observed that hydrodynamic dispersion causes solutes to move both faster and slower than the advective front, resulting in smooth concentration profiles.

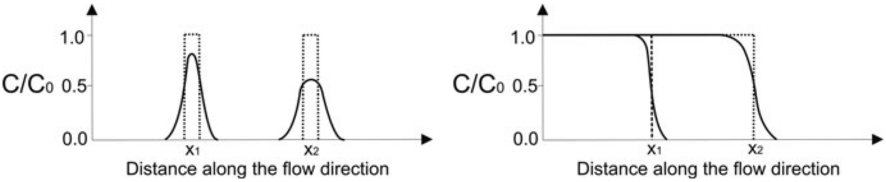


Fig. 19.4 Relative concentration as a function of distance in the flow direction, illustrating the combined effect of advection and hydrodynamic dispersion on the migration of a contaminant after an instantaneous (*left*) and a continuous release (*right*)

Since hydrodynamic dispersion strongly influences the migration of contaminants in groundwater, the determination of dispersion coefficients and in particular of hydromechanical dispersivities turns out to be of key importance for the description of contaminant transport. Research studies conducted primary on longitudinal dispersion showed a remarkable difference between laboratory (typically 10^{-4}–10^{-2} m) and field scale dispersivities, which were often found to be orders of magnitude larger. This has led to extensive research on the role of field scale macro-dispersion, e.g. by applying stochastic approaches. The reader interested in stochastic approaches to describe subsurface flow and transport can refer to the work of Dagan (1989) and Gelhar (1993), among others.

19.2.4 Sorption

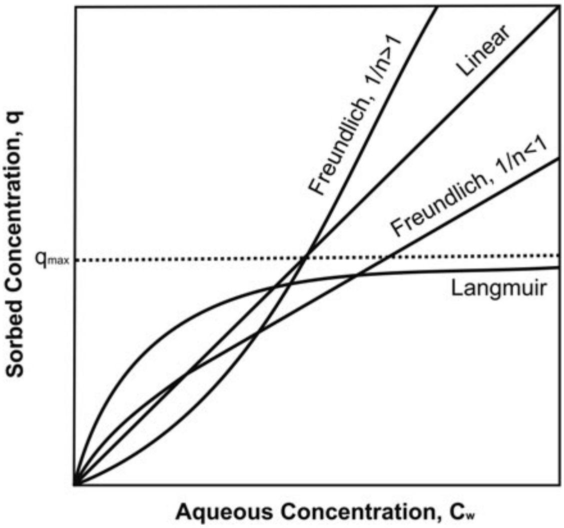
In aquifer sediments, as in the upper (unsaturated) soil layers, contaminants are temporarily removed from the groundwater by interaction with the solid matrix by chemical, physical or electrostatic forces. This process is generally called sorption. Two sorption phenomena can be typically distinguished: adsorption and absorption. Adsorption refers to processes in which the contaminant accumulates on the surface of a soil particle. Absorption describes processes in which the contaminant penetrates into a separate phase e.g.organic matter in soils. In heterogeneous soils and aquifers the two processes may occur simultaneously. For more comprehensive information see the review by Allen-King et al. (2002).

The affinity of a dissolved contaminant for sorption by aquifer sediments is commonly described by sorption isotherms (Grathwohl 1998; Schwarzenbach et al. 1993). In the simplest case, the concentration in the solid, C_s [M M^{-1}], is proportional to the equilibrium contaminant concentration in the aqueous phase $\mathrm{C_w}$ [M L^{-3}]:

$$C_s = K_d C_w \tag{19.12}$$

where K_d [L^3 M^{-1}] is the distribution coefficient. K_d is the ratio between the contaminant concentration in solid soil particles and groundwater; it represents the slope of the linear sorption isotherm (see Fig. 19.5).

Fig. 19.5 Schematic illustration of sorption isotherms



However, especially if adsorption prevails, the concentration in soil usually depends non-linearly on the concentration in the groundwater (non-linear sorption). For non-linear isotherms (Fig. 19.5) the Freundlich and Langmuir models are often used. The Freundlich isotherm can be written as:

$$C_s = K_{Fr} C_w^{1/n} \tag{19.13}$$

where K_{Fr} [(M M^{-1})/(M L^{-3})$^{1/n}$] is the Freundlich sorption coefficient and *1/n* is an empirical exponent. The Freundlich isotherm reduces to the linear model, analogous to Eq. (19.12), if $1/n = 1$.

The Langmuir model represents another non-linear isotherm which takes into account a maximum sorption capacity:

$$C_s = \frac{K_L C_{s,\max} C_w}{1 + K_L C_w} \tag{19.14}$$

where $C_{s,\max}$[M M^{-1}] is the maximum concentration in the solid and K_L [L^3 M^{-1}] is the Langmuir sorption coefficient. For $K_L C_W << 1$ the Langmuir isotherm predicts a linear relationship analogous to Eq. (19.12).

In groundwater, sorption causes retardation of the advective and dispersive transport of dissolved contaminants. Sorption processes typically influence the time it takes for a contaminant to travel a certain distance and the overall plume to reach steady-state. After steady state is achieved, the sorption capacity generally does not influence the length of the plume (Liedl et al. 2005). However, sorption processes play an important role under transient flow and transport conditions (Cirpka 2005; Prommer et al. 2002).

Different approaches can be followed to describe sorption in the framework of a contaminant reactive transport model (see Section 19.3). If the characteristic time for sorption processes is considerably shorter than the characteristic time scale of transport, it is appropriate to consider sorption at equilibrium (local equilibrium approach) and to apply one of the isotherm models described above (Fig. 19.5). In contrast, when the local equilibrium assumption is not valid, then sorption must be described as a rate-limited reaction process (e.g. Barry et al. 2002; Cirpka 2005).

19.2.5 Biodegradation

Although abiotic processes can contribute to contaminants degradation, biological degradation is by far the most significant mass removal process of organic contaminants in groundwater.

In the case of organic contaminants, the biologically catalyzed degradation frequently, although not necessarily, leads to the conversion of much of the carbon, nitrogen, phosphorus, sulfur and other elements present in the original compound to inorganic products (mineralization). Microorganisms are present at large quantities in groundwater systems, although numbers are typically less than in the unsaturated upper layer (see Chapter 13 by Swartjes et al., this book). They use naturally occurring and many synthetic organic contaminants for their growth. From an environmental perspective, a crucial role of microorganisms is the degradation of contaminants and, in particular, their ability to bring about detoxification, the transformation of a contaminant molecule into a less harmful product (Alexander 1998). Microbially mediated degradation of contaminants can follow a wide variety of metabolic pathways. Therefore, under different environmental conditions, processes such as aerobic oxidation, anaerobic oxidation, anaerobic reduction, fermentation, cometabolism, et cetera can result in the degradation of diverse organic contaminants.

Although significant progress has been made in the recent years in the study of subsurface microbial activities, the details of many degradation pathways are still unknown. However, it is possible to identify some basic requirements of all biodegradation processes (Cookson 1995):

- the presence of microorganisms with the capability to degrade the target contaminant(s);
- the presence of a substrate that can be used as an energy and carbon source;
- the presence of an appropriate electron acceptor (e.g. O_2, NO_3^-, Fe^{3+}, SO_4^{2-}, et cetera);
- the presence of optimal environmental conditions (e.g. moisture, pH, temperature) adequate for the enzymatically catalyzed reactions;
- the presence of nutrients necessary to support the microbial cell growth and enzyme production;
- the absence of toxic substances to the microorganisms.

Shallow aquifer systems, where the vast majority (80–90%, Chapelle 2001) of documented cases of groundwater contamination occurs, are not sterile environments. On the contrary, they all contain viable consortiums of microbial species in a more or less active state. It is unlikely that one specific microorganism is sufficient for successful remediation of a contaminated site; in fact, complete mineralization is often the result of the activity of mutually dependent microbial communities: a degradation sequence occurs where a second organism degrades the metabolic products of the first, and a third will use the products of the second et cetera.

Microorganisms affect not only the fate of contaminants, but also geochemical conditions in the groundwater. Many important chemical species in groundwater systems, such as oxygen, organic and inorganic carbon species, sulfur species, nitrogen species and iron and manganese species, move through microbial food chains. Thus, microbial activities influence and, at the same time, are strongly influenced by the subsurface geochemical conditions. The understanding of the nature of biogeochemical cycles is fundamental to model fate and transport of chemical species in subsurface systems. For more details on degradation of contaminants (Natural Attenuation), see Chapter 22 by Peter et al., this book.

Chemically, the degradation of the organic contaminants involves redox transformations. Through complex electron transfer chains, heterotrophic microorganisms are able to transfer the electrons extracted from the oxidation of an organic contaminant to terminal electron acceptors present in the surrounding environment. Thus, the electron acceptors are reduced and new chemical entities such as reduced dissolved, adsorbed and solid species are produced. In a contaminated aquifer, microorganisms degrade the organic contaminants while reducing different electron acceptors naturally present in groundwater such as: dissolved oxygen, nitrate, iron oxides and hydroxides, sulfate and carbon dioxide. Considering toluene as an example and assuming steady state conditions for microbial biomass, biodegradation reactions through different TEAPs (Terminal Electron Acceptor Processes) are listed in Table 19.1. The reactions are characterized by different energetic yields that determine a sequential order of electron acceptor consumption.

Biodegradation of toluene through different TEAPs has been modeled in a batch system with a double Monod kinetic expression including inhibition terms to represent the sequential order of the degradation processes (Rolle et al. 2008a). The results of this illustrative example, developed in the MATLAB® computing environment, are shown in Fig. 19.6. The TEAPs have been implemented in the

Table 19.1 Processes and corresponding redox reactions for toluene degradation

Process	Redox reaction
Aerobic respiration	$C_7H_8 + 9O_2 \rightarrow 7CO_2 + 4H_2O$
Denitrification	$C_7H_8 + 7.2NO_3^- + 7.2H^+ \rightarrow 7CO_2 + 3.5N_2 + 7.6H_2O$
Manganese reduction	$C_7H_8 + 18MnO_2 + 36H^+ \rightarrow 7CO_2 + 18Mn^{2+} + 22H_2O$
Iron reduction	$C_7H_8 + 36Fe(OH)_3 + 72H^+ \rightarrow 7CO_2 + 36Fe^{2+} + 94H_2O$
Sulfate reduction	$C_7H_8 + 4.5SO_4^{2-} + 4.5H^+ \rightarrow 7CO_2 + 4.5HS^- + 4H_2O$
Methanogenesis	$C_7H_8 + 5H_2O \rightarrow 4.5CH_4 + 2.5CO_2$

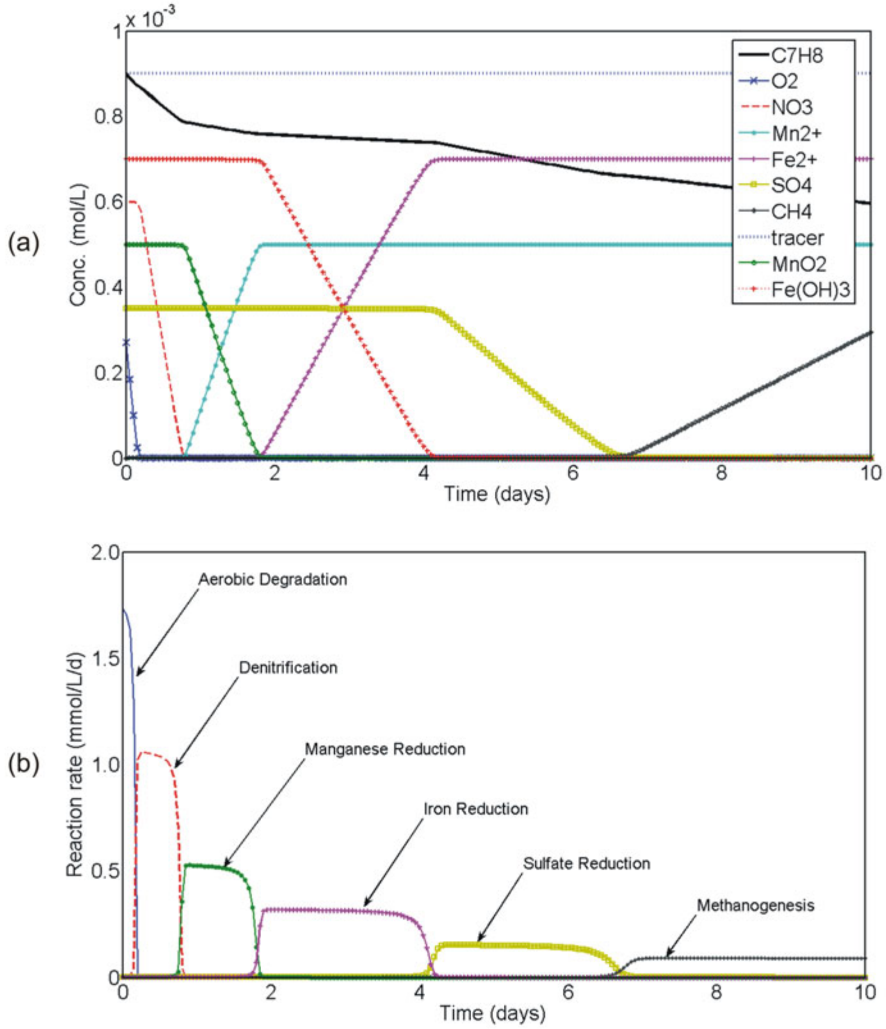


Fig. 19.6 Sequential degradation of toluene through different TEAPs: (**a**) temporal concentration profiles of electron donor, electron acceptors, metabolic by-products and a conservative tracer; (**b**) computed reaction rates of the different TEAPs determining a temporal redox zonation

model strictly following the energetic yield of the redox reactions allowing very little overlapping between the different processes. The decreasing trend in the toluene concentration results from biodegradation with the sequential utilization of the different electron acceptors (Fig. 19.6a). Dissolved oxygen and nitrate are consumed first, followed by the solid species MnO_2 and $Fe(OH)_3$ producing dissolved manganese and iron. Subsequently, sulfate starts to be depleted and finally methanogenesis takes place. The sequential order of the TEAPs can be effectively visualized by plotting the transient reaction rates of the different processes (Fig. 19.6b).

In groundwater systems the biodegradation potential not only depends on the relative energy yields of the different reactions, but also on the availability of organic substrates, electron acceptors, and active microbial populations. Hence, not all possible processes are necessarily occurring in a given system. Most pristine shallow groundwater systems are oligotrophic (organic-poor) and hence may remain aerobic, if the dissolved oxygen is not consumed (Chapelle 2001). For aquifers which become contaminated with oxidizable organic contaminants, or which have naturally high concentrations of dissolved organic matter in their recharge waters or solid matrix, dissolved oxygen may become depleted, thereby allowing the emergence of anaerobic processes. In contaminated aquifers, the differences in the distribution of microbial activity lead to physical separation (spatial and temporal) of the zones in which particular TEAPs dominate. Thus, the sequential utilization of electron acceptors by microorganisms gives rise to a characteristic redox zonation, frequently observed down-gradient of contaminated sites (Baedecker and Back 1979; Christensen et al. 2000; Brun et al. 2002; Chapelle et al. 1995; Rolle et al. 2008a). The spatial distribution of the TEAPs allows to recognize two different patterns of biodegradation in contaminant plumes: fringe and core processes. Highly bioactive zones develop at the plume fringe, where the oxidizable organic contaminants are brought into contact with electron acceptors (e.g. dissolved oxygen, nitrate and sulfate) through diffusion/dispersion controlled mixing processes. Core processes, such as manganese reduction, iron reduction and methanogenesis, take place inside the contaminant plume.

The identification of different redox zones (as shown in the previous example for the oxidation of toluene) is of fundamental importance to assess the overall potential for degradation of most organic contaminants in groundwater. In fact, attenuation of specific organic contaminants is strongly influenced by redox zonation. This concept is valid not only for contaminants degrading preferentially (and/or most rapidly) in an oxidizing environment (e.g. hydrocarbons such as BTEX and PAHs), but also for contaminants (e.g. chlorinated solvents) that are typically reduced under anaerobic conditions (Wiedemeier et al. 1999). For the latter group of contaminants, the prevailing redox environment, determined by the competition for electron donors among different members of the microbial community, strongly influence the efficiency of degradation.

19.3 Contaminant Transport Models

Mathematical models for contaminant transport have become increasingly important tools to assess the risk due to contaminant migration and to support the design of Risk Management solutions, such as groundwater remediation. They provide a rational framework to rigorously incorporate a wide variety of processes, either physical or biogeochemical. Therefore, they offer an ideal playground to integrate and test process knowledge gained from theoretical, experimental (laboratory) and field investigations.

The development and application of contaminant transport models are motivated by fundamental questions such as the need to assess the impact of the contamination on environmentally important receptors located down-gradient of the source zone, to estimate time-scales and optimal design of a remediation scheme and its sensitivity to changes in physical or biogeochemical conditions (Prommer and Barry 2005). The results of transport models always bear a degree of uncertainty that originates from the usually sparse dataset available and from the incomplete hydrogeological and hydrogeochemical site characterization, lack in process understanding and parameter ambiguity due to spatial heterogeneity. Nonetheless, transport models are characterized by high flexibility and allow a quick investigation of a number of possible scenarios. Therefore, they remain an indispensable tool to gain an improved understanding of factors controlling the contaminant fate and transport.

19.3.1 Governing Equations

Because of the impossibility of mathematically describing the complicated geometry of the solid surfaces that delimit the flow domain in a natural porous medium, the vast majority of groundwater flow and transport models are based on the continuum approach (Bear 1972). According to this mathematical approach, the actual multiphase porous medium is substituted with a fictitious continuum, a structureless substance to any point of which kinetic and dynamic variables and parameters can be assigned. These properties, averaged over a representative elementary volume (REV), are a continuous function of the spatial coordinates of a point and of time and allow the description of flow and transport phenomena in porous media by means of partial differential equations (PDEs). The selected REV should be much larger than the microscopic scale of heterogeneity associated with the presence of solid particles and pore spaces in the porous medium and much smaller than the considered domain.

The governing groundwater flow equation can be derived by applying the mass conservation principle to an aquifer control volume (Bear 1972, 1979; Zheng and Bennet 2002) and can be written as:

$$\frac{\partial}{\partial x_i}\left(K_{ij}\frac{\partial h}{\partial x_j}\right)+q_s=S_s\frac{\partial h}{\partial t} \tag{19.15}$$

where:

$x_{i,j}$ = distance along the respective Cartesian coordinate axis [L];
t = time [T];
K_{ij} = tensor of hydraulic conductivity [L T^{-1}];
h = hydraulic head [L];
q_s = fluid source/sink term [$\mathrm{L}^3\mathrm{L}^{-3}\mathrm{T}^{-1}$];
S_s = specific storage [$\mathrm{L}^3\mathrm{L}^{-3}\mathrm{L}^{-1}$];

Appropriate initial and boundary conditions, geometry of the model domain and specified values of the physical coefficients together with the partial differential equation for flow (Eq. (19.15)) define an individual flow problem.

In a similar way the governing transport equation for a dissolved contaminant can be derived based on the mass conservation of the contaminant through an aquifer control volume. The partial differential equation describing the fate and transport of a dissolved contaminant in a three-dimensional groundwater flow system can be written as follows (Zheng and Bennet 2002):

$$\frac{\partial(nC)}{\partial t} = \frac{\partial}{\partial x_i}\left(nD_{ij}\frac{\partial C}{\partial x_j}\right) - \frac{\partial}{\partial x_i}(nv_iC) + q_sC_s + \sum R_n \tag{19.16}$$

where:

n = porosity of the aquifer sediments [–];
C = contaminant concentration in groundwater [M L^{-3}];
t = time [T];
$x_{i,j}$ = distance along the respective Cartesian coordinate axis [L];
D_{ij} = hydrodynamic dispersion coefficient tensor [L^{2}T^{-1}];
v_i = average linear velocity [L T^{-1}];
q_s = volumetric flow rate per unit volume of aquifer representing fluid sources (positive) and sinks (negative) [T^{-1}];
C_s = concentration of the source or sink flux [M L^{-3}];
ΣR_n = chemical reaction term [M L^{-3}T^{-1}].

The chemical reaction term added to the general advective-dispersive equation allows us to describe the principal biogeochemical processes (e.g. sorption, biodegradation, precipitation/dissolution, et cetera) taking place in the subsurface.

The mathematical problem of contaminant transport, expressed in terms of governing equations, initial and boundary conditions, together with the applicable flow and transport parameters, and the information on sources and sinks and on the reaction term, can be solved to obtain the concentration distribution in the modeled region at a time of interest. Alternatively, the concentration breakthrough in time can be calculated at a location of interest.

19.3.2 Mathematical Models

There are two main procedures for obtaining the solution of a mathematical model: analytical and numerical methods. These methods result in two different types of models used to solve groundwater flow and contaminant transport problems. Generally, analytical models can be applied only under a number of simplifying assumptions, whereas more complex groundwater flow and contaminant transport problems require numerical solutions. Most of the practical applications at contaminated sites need numerical solutions of the groundwater flow and/or contaminant transport equations. Therefore, some sort of spatial and temporal discretization of the problem governing equations has to be performed. The result of the discretization is the approximation of the governing partial differential equations with a set of algebraic equations typically solved by a computer code.

The often cited numerical codes such as MODFLOW (McDonald and Harbaugh 1988) and FEMWATER (Yeh and Ward 1980) are used to solve the governing three-dimensional groundwater flow equation (Eq. (19.15)) and to calculate the distribution of hydraulic heads which represent a prerequisite for the computation of the velocity field necessary for the solution of the contaminant transport equation.

The numerical solution of contaminant transport equations is an area of active research. An exhaustive discussion of numerical methods, however, is beyond the scope of this book chapter. In the following a brief overview of principal numerical techniques and strategies is presented. For further details the interested reader can refer to the literature (e.g. Barry et al. 2002; Bear 1979; Steefel and MacQuarrie 1996; Zheng and Bennet 2002; Zheng and Wang 1999, among others).

Most numerical methods for solving the advection-dispersion equation can be classified as Eulerian, Lagrangian or mixed Eulerian-Lagrangian. In Eulerian methods, the transport equation is solved in a fixed spatial grid. Finite difference, finite element and finite volume methods are primary examples for this class of solution methods. Eulerian methods are generally mass conservative and handle dispersion-dominated problems both accurately and efficiently. However for advection-dominated problems, encountered in many field situations, an Eulerian method may be susceptible to excessive numerical errors (Zheng and Wang 1999). A typical error is “numerical dispersion”, which indicates the artificial, non-physical dispersion resulting from the numerical approximation associated with the discretization of the model domain. In order to overcome these problems and minimize numerical errors, restrictively small grid spacing and time-steps may be required. Alternatively, higher order finite difference (or finite element) methods such as the total-variation-diminishing (TVD) method can be used.

In the Lagrangian approach, the transport equation is solved in either a deforming grid or a deforming coordinate in a fixed grid through particle tracking. This approach provides a highly efficient solution to advection-dominated problems virtually free of numerical dispersion. However, problems such as local mass balance errors and numerical instabilities may arise.

Mixed Eulerian-Lagrangian methods (e.g. the widely used methods of characteristics) attempt to combine the advantages of both approaches by solving the advection term with a Lagrangian method (particle tracking) and the dispersion term and the other terms with an Eulerian approach (finite difference or finite elements). Concerning the solution of coupled reactive transport problems, two different strategies can be distinguished:

- one-step or global implicit approach, which solves the governing transport equations, including transport and reaction terms, simultaneously. An example of contaminant reactive transport model using this method is MIN3P (Mayer et al. 2002);

- sequential (iterative or non-iterative) approach (operator splitting), which solves separately the transport and reaction terms. Widely used multispecies and multicomponent reactive transport models such as RT3D (Clement 1997) and PHT3D (Prommer et al. 2003) are based on this method.

19.3.3 Model Application

The application of a groundwater flow and transport model, incorporating the principal hydrogeological and hydrogeochemical conditions of a particular contaminated site, can be described as a succession of several steps (e.g. Bear et al. 1992; Zheng and Bennet 2002):

(1) Formulation of objectives: the identification of the goals and purposes of the modeling study, which could include an improvement of the understanding of the transport regime at a particular site, or setting the priority for remediation depending on contaminant transport characteristics.
(2) Review and interpretation of the available data: compilation of all the sources of information from the site characterization.
(3) Development (or improvement) of the conceptual model based on the compiled information including geological, hydrogeological, geochemical data, together with information on the site history and presumable location of contaminant sources.
(4) Field data collection: at this stage the need to collect new data in addition to those available from the site characterization should be evaluated.
(5) Selection of an appropriate computer code: the choice should depend on different factors including not only the capability to represent the principal processes determining the contaminant transport at the site, but also on the goals of the modeling study and the expertise of the modeler.
(6) Construction of the contaminant transport model: by preparing the input data, the conceptual model is translated in the site-specific mathematical model.
(7) Model calibration and sensitivity analysis: after the model has been run with some initial estimates of model parameters, these input parameters must be adjusted, either manually (trial-and-error) or assisted by a computer optimization code, in order to match the field observations to a reasonable degree. Sensitivity analysis is performed before and after calibration to test the sensitivity of the computed model results to some input parameters (e.g. Hill and Tiedeman 2007).
(8) Predictive simulations: after the contaminant transport model has been calibrated it can be run to predict future contamination scenarios and/or the impact of remediation activities.
(9) Uncertainty analysis: many uncertainty sources are associated with a field scale modeling study such as incomplete understanding of physical and biogeochemical processes, spatial and temporal variability of model parameters and/or

boundary and initial conditions, errors associated with field measurements, et cetera, and these are increased by the typical availability of only sparse spatial and temporal field datasets. Different methods have been proposed for uncertainty analysis in reactive transport models including Monte Carlo simulations and stochastic models (e.g. Hill and Tiedeman 2007; Zheng and Bennet 2002).

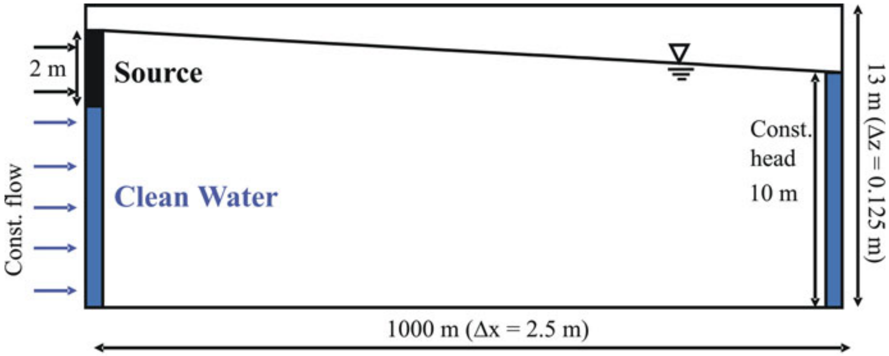
It is important to note that the short list of guidelines for model application given above should not be regarded as a merely sequential procedure, but rather as an iterative process where the outcomes of a given action are made available for preceding or subsequent steps. This allows a continuous improvement of the understanding of the contaminant fate and transport at a specific site. Thereby, one of the most important characteristics of a contaminant transport model – its capability of providing a quantitative framework to integrate site-specific information and to test and refine the site conceptual model – will be fully exploited.

In the following sections, practical applications of contaminant reactive transport modeling are illustrated for a typical petroleum hydrocarbon contamination scenario and for a field study at a landfill site, where ammonium was the principal contaminant of concern.

19.4 Reactive Transport Scenarios

Several two-dimensional reactive transport scenarios were simulated with the aim of evaluating the influence of different parameters on contaminant fate and transport. To this purpose, the fate of a toluene plume in a shallow unconfined aquifer was investigated. The influence of different groundwater flow, contaminant transport and kinetic parameters on plume migration was assessed through a sensitivity analysis.

In Fig. 19.7 the geometry and the conceptual model of toluene transport in groundwater is shown. As depicted in this figure, the modeled aquifer is considered a two-dimensional unconfined system with groundwater entering the domain from the left boundary and flowing, with a dominant horizontal component, towards the right. Toluene dissolving from a LNAPL source in the upper groundwater layer was considered as model contaminant. As shown above (Table 19.1), the release of an oxidizable organic contaminant, such as toluene, in a pristine aquifer can activate multiple biogeochemical processes (Chapelle 2001; Hunter et al. 1998). For the sake of simplicity the attention is focused exclusively on the aerobic degradation of toluene, in this modeling example. As the toluene plume migrates down-gradient, a reactive fringe forms in particular at the lower edge of the plume. At the fringe the two reactants, i.e. the oxidizable contaminant (electron donor) and oxygen (electron acceptor) mix, allowing aerobic degradation of toluene to proceed. As highlighted by previous modeling studies (e.g. Cirpka et al. 1999; Ham et al. 2004; Liedl et al. 2005; Maier and Grathwohl 2006; Rolle et al. 2005), under steady state conditions, the principal process controlling the mixing of reactants is transverse dispersion. The presence of mixing controlled enhanced biodegradation activity at the plume



Flow & Transport
Steady state flow
$K = 43.2$ m/d
$v = 1$ m/d
$n_e = 10\%$
$\alpha_L = 0.02$ m
$\alpha_T = 0.002$ m

Concentrations

Source: Toluene = 0.01mM
$O_2 = 0$ mM
$Na^+ = Cl^- = 1$ mM

Groundwater: Left BC, IC Toluene = 0 mM
$O_2 = 0.4$ mM
$Na^+ = Cl^- = 1$ mM
pH = 7
pe = 13.95

Fig. 19.7 Geometry, conceptual model and parameters for an example of toluene transport in groundwater

fringe have been identified in some field studies (e.g. Anneser et al. 2008; Lerner et al. 2000) and, recently, demonstrated in a series of laboratory experiments (Bauer et al. 2008, 2009a, b; Rolle et al. 2008b, 2009).

Figure 19.7 also reports the principal parameters used in the simulations. The extent of the model domain was chosen to represent a typical contamination scenario for a shallow unconfined aquifer. To evaluate a suitable spatial discretization for the model, a range of different grid resolutions was tested. One of the major concerns was that artificial mixing due to numerical dispersion would cause an overestimation of the fringe biodegradation reaction (Cirpka et al. 1999). Therefore, a fine vertical grid discretization ($\Delta z = 0.125$ m) and the HMOC solver for advection (hybrid method of characteristics, Zheng and Wang (1999)) were selected to limit numerical dispersion and to be able to accurately capture the thickness of the reactive fringe zone. The flow simulations were carried out for steady-state conditions using the numerical code MODFLOW (McDonald and Harbaugh 1988). A constant hydraulic head boundary condition was used at the downstream boundary and a constant groundwater flow condition was used at the upstream boundary.

The numerical code selected for the transport simulations was PHT3D (Prommer et al. 2003). It combines the transport simulator MT3DMS (Zheng and Wang 1999) with the geochemical model PHREEQC-2 (Parkhurst and Appelo 1999) for the computation of coupled transport and reactive processes. As initial condition for the transport simulation, it was assumed that the aquifer was uncontaminated. The contaminant originates from the upper-left boundary with a constant concentration of

1×10^{-5} mol/L; an underlying flow of clean water with simplified water chemistry and a dissolved oxygen concentration of 4×10^{-4} mol/L were selected as boundary conditions. The simulations were run for a total simulation time of 1200 days, subdivided into 600 time steps.

In steady state groundwater plumes, a balance between microbial growth and decay can be considered and the aerobic degradation of toluene can be summarized by the following redox reaction (Wiedemeier et al. 1999):

$$\mathrm{C_7H_8} + 9\mathrm{O_2} + 3\mathrm{H_2O} \rightarrow 7\mathrm{HCO_3^-} + 7\mathrm{H^+} \tag{19.17}$$

For the present modeling study a so-called double Monod formulation was used to describe the reaction kinetics. It is a commonly used expression (e.g. Barry et al. 2002) and it takes into account the reaction rate dependence on both the electron donor (toluene) and the electron acceptor (oxygen), as follows:

$$\frac{\partial [C_7H_8]}{\partial t} = -k \frac{[C_7H_8]}{K_{C_7H_8} + [C_7H_8]} \frac{[O_2]}{K_{O_2} + [O_2]} \tag{19.18}$$

where k is the kinetic rate constant [mol L^{-1} s^{-1}] and K_{C7H8}, K_{O2} are the half-saturation Monod constants for toluene and oxygen [mol L^{-1}]. A value of 5×10^{-11} mol L^{-1} s^{-1} was selected for the kinetic constant in the base modeling scenario while the half-saturation constants were set to 1×10^{-5} mol L^{-1} for all the simulation runs. The results of the base scenario, after a total simulation time of 1200 days, are shown in Fig. 19.8a. It can be observed that despite the continuous supply of toluene, the plume reaches a steady state length $L = 712.5$ m, when the contaminant mass flux removed by aerobic degradation equals the contaminant input. Evolving from this base case a number of additional scenarios was computed to identify and quantify the influence and sensitivity of individual parameters on steady state plume length.

Figure 19.9, for example, shows the effect of transverse dispersivity. This parameter largely controls the mixing of reaction partners and, therefore, it is of pivotal importance for microbially mediated fringe degradation reactions. As shown in Fig. 19.9 (squares), longer steady state plumes were computed with decreasing transverse dispersivity (α_T). A linear relationship between the plume length and the inverse of the transverse dispersivity was derived also in similar two-dimensional studies with different boundary conditions compared to the present scenario. For instance, Liedl et al. (2005) derived the following analytical solution of steady state plume length in a completely contaminated aquifer with supply of the electron acceptor just from the top, through the water table:

$$L = \frac{4M^2}{\pi^2 \alpha_T} \ln \left(\frac{4}{\pi} \frac{\gamma c_D^0 + c_A^0}{c_A^0} \right) \tag{19.19}$$

where $c_D{}^0$ and $c_A{}^0$ are the initial concentrations of electron donor and acceptor, α_T is the vertical transverse dispersivity, M is the source thickness and γ is the

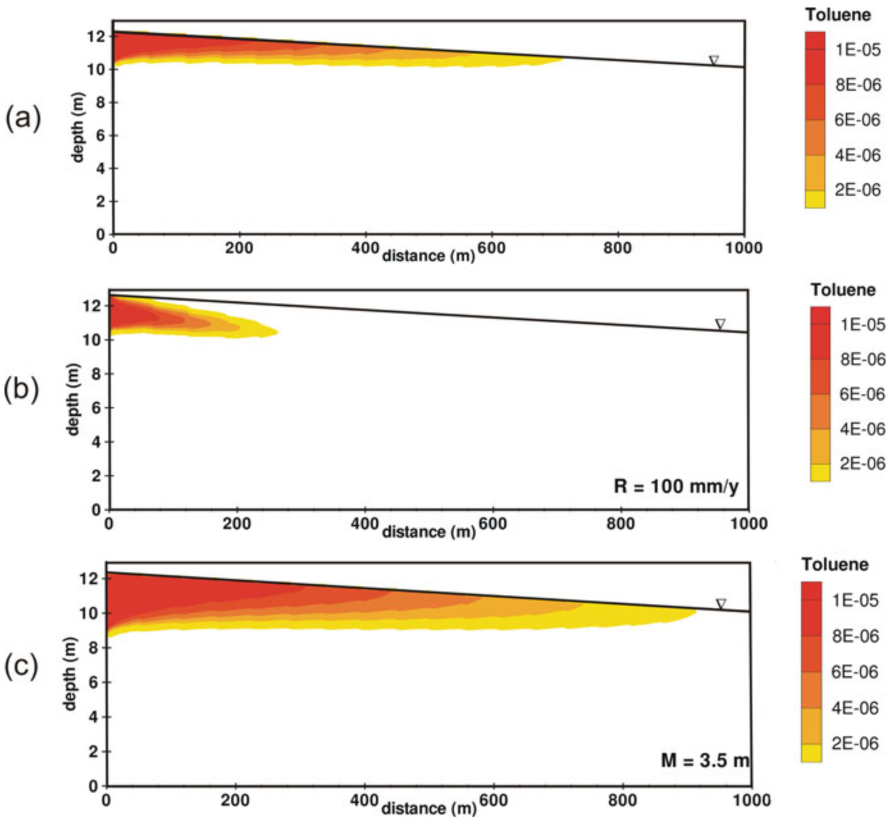


Fig. 19.8 Simulated steady state toluene plumes: (**a**) base scenario; (**b**) base scenario with a recharge of 100 mm year^{-1}; (**c**) plume from a thicker (3.5 m) contamination source

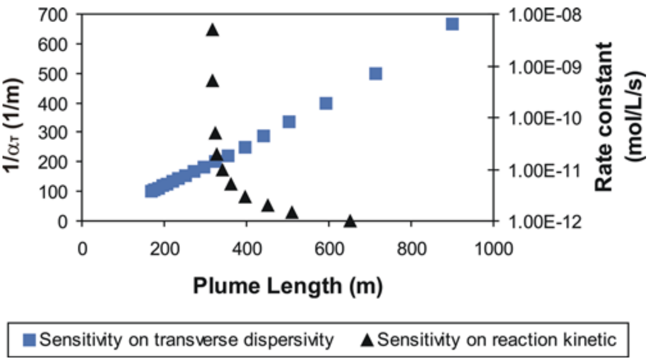


Fig. 19.9 Sensitivity of simulated steady state plume length to transverse dispersivity and to biodegradation rate constants

stoichiometric ratio (number of moles of acceptor needed to degrade 1 mole of donor). Maier and Grathwohl (2006) report a similar empirical correlation:

$$L = 0.5\frac{M^2}{\alpha_T}\left(\frac{C_D\gamma}{C_A}\right)^{0.3} \tag{19.20}$$

In the same plot the dependence of the plume length on biodegradation rate constant is shown (triangles). The reaction kinetics have a strong influence on steady state plume length as long as the reaction is slow and acts as limiting factor (lower portion of the plot). When the degradation rate is fast, the supply rate of electron acceptor to the plume by transverse mixing becomes the limiting process and the plume length turns out to be independent of the biodegradation rates. Biodegradation kinetics also determines the shape of the contaminant plume and the consumption of electron acceptors, as can be observed from the results of the simulations shown in Fig. 19.10. Simulations of toluene transport and aerobic degradation were performed using two different degradation rate constants. With fast, almost instantaneous degradation kinetic, the steady state plume is shorter and the vertical gradients of electron donor and acceptor do not overlap. When the degradation rate constant was set to a lower value, a longer plume was computed as well as the overlapping of the toluene and oxygen vertical gradients. This overlapping indicates that, in the second case, the degradation rate and no more the transverse mixing is the limiting factor for the overall biodegradation.

Also the amount and quality of the infiltrating water from the surface affects the fate and transport of contaminant plumes in shallow groundwater systems. The recharge enhances the dilution of the plume and acts as a new source of electron acceptors and nutrients, which adds a positive effect on the overall biodegradation capacity of the aquifer. Simulations were performed with the composition of the

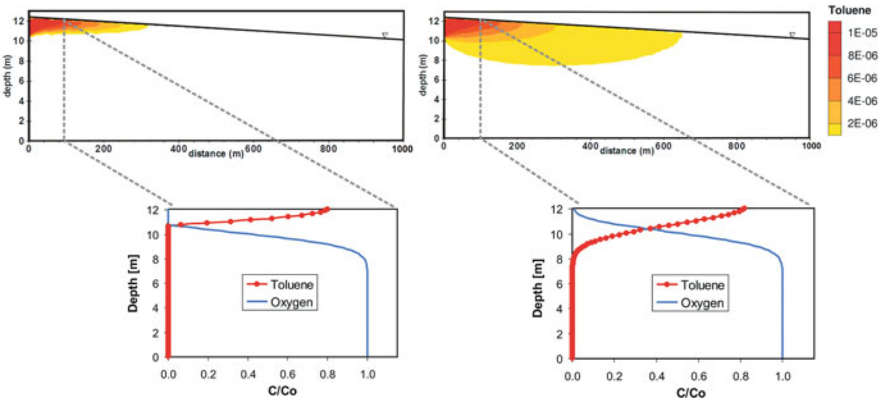


Fig. 19.10 Simulated toluene plumes assuming a fast ($k = 5\times10^{-10}$ mol L^{-1} s^{-1}) and a slow ($k = 1\times10^{-12}$ mol L^{-1} s^{-1}) biodegradation rate constant and corresponding vertical profiles of toluene and oxygen at a longitudinal distance of 100 m

infiltrating recharge water assumed to be identical to the clean water coming into the aquifer from the left boundary of the model domain. The result of one simulated scenario with 100 mm/year recharge is shown in Fig. 19.8b. In comparison to the base scenario, the effect of recharge determines a smaller toluene plume due to the establishment of a second reactive fringe on the top of the plume where the infiltrating water comes into contact with the dissolved contaminant. This second reactive fringe also causes a deeper infiltration of the toluene plume in the aquifer.

As shown in the analytical formula and empirical correlation reported above (Eqs. (19.19) and (19.20)), the steady state plume length also depends on the thickness of the contamination source. An increase of the source thickness results in higher mass fluxes in the aquifer and thus in longer steady state plumes as can be seen in Fig 19.8c where a source thickness of 3.5 m was used instead of the 2 m thickness of the base scenario.

Simulations have been performed starting from the base scenario using different source thicknesses in a range of 1.5–3.5 m and with transverse dispersivity values of 2×10^{-3} and 5×10^{-3} m. The results confirmed the dependency of the steady-state plume length on the square of the source thickness as can be observed in Fig. 19.11. This plot shows the steady state plume length as a function of source thickness for different values of transverse dispersivity, at a double-logarithmic scale.

The simulations performed above assumed a homogeneous aquifer system. However, aquifer heterogeneity, in particular physical heterogeneity, characterizes many contaminated sites and strongly influences contaminant plumes migration. A considerable research effort has been dedicated on including the complex heterogeneity of many natural porous media into either deterministic or stochastic contaminant transport models. For the present reactive transport modeling study, just the effect of flow focusing in well-defined high permeability inclusions was considered. The focusing of flow in high-permeability zones in heterogeneous porous media causes the streamlines to converge and diverge depending on the permeability

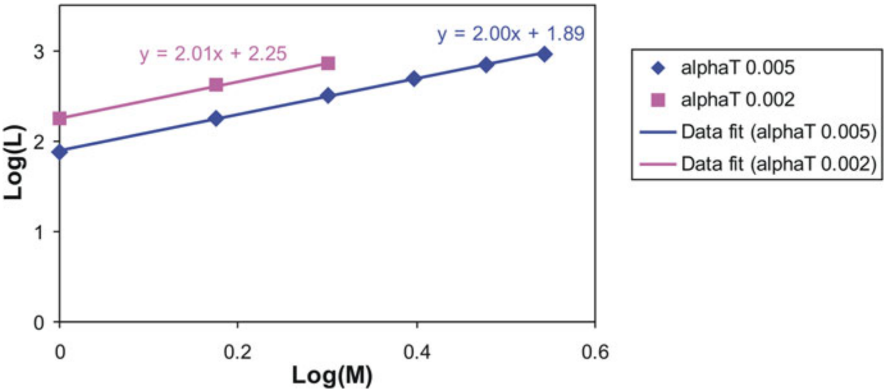


Fig. 19.11 Steady state plume length as a function of source thickness for different values of transverse dispersivity

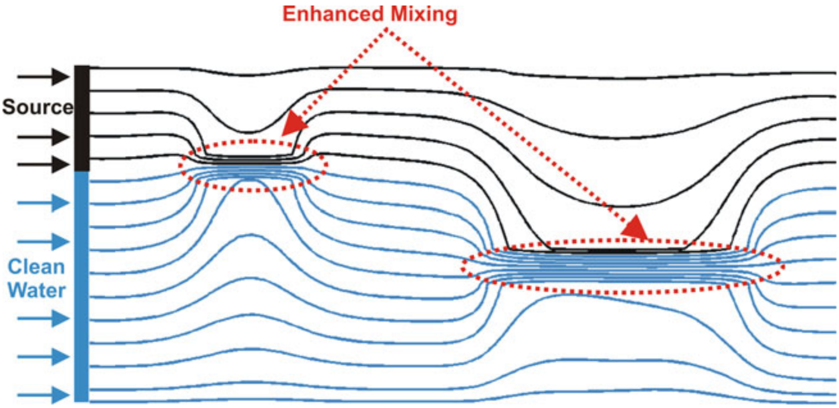


Fig. 19.12 Focusing of pathlines into high permeability zones included into the simulation domain. Pathlines computed using MODPATH (Pollock 1994)

distribution (Werth et al. 2006). When flow is focused in a high-permeability zone the velocity increases and the distance required for a solute to cross a given number of flow lines decreases. This focusing results in an enhancement of transverse mixing and, therefore, of reactions occurring at the plume fringe (Fig. 19.12). These concepts were experimentally demonstrated at the laboratory scale, in tank systems with both instantaneous abiotic (Rolle et al. 2008c, 2009) and microbially mediated reactions (Bauer et al. 2009a).

In the present scenario modeling a rectangular high-permeability lens, 175 m long and 0.5 m thick, was included in the simulation domain. Different model runs were performed using different hydraulic conductivity ratios between the high-permeability lens (K_L) and the surrounding aquifer sediments (K). Also, at this larger scale, the focusing of the reactive fringe caused a significant increase of mixing and reaction of toluene and oxygen. For example, a ratio of hydraulic conductivities (K_L/K) of 4 was sufficient to result in a steady state plume approximately 100 m shorter with respect to the base case steady state plume length (Fig. 19.8a). This example conceptually illustrates the effect of mixing and reaction enhancement due to flow focusing in a simplified aquifer geometry and hydraulic conductivity field at the typical scale of groundwater contamination. Under natural conditions, often characterized by a high degree of physical heterogeneity, a high variability of the mixing and reaction enhancement is likely to be expected. As shown by Werth et al. (2006), this variability reflects the extent to which the reactive fringe of the plume is focused into different high-permeability inclusions and entails a considerable variability in the prediction of reactive transport (Cirpka et al. 2008).

Besides the flow focusing in high permeability formations embedded into a less permeable porous medium (e.g. coarse sand and gravel lenses and layers in a fine sand aquifer), other hydrogeological configurations and/or stresses such as the presence of impervious bodies (e.g. constructions), the pumping from a partially penetrating well, et cetera, can result in focusing of groundwater flow. In Fig. 19.13,

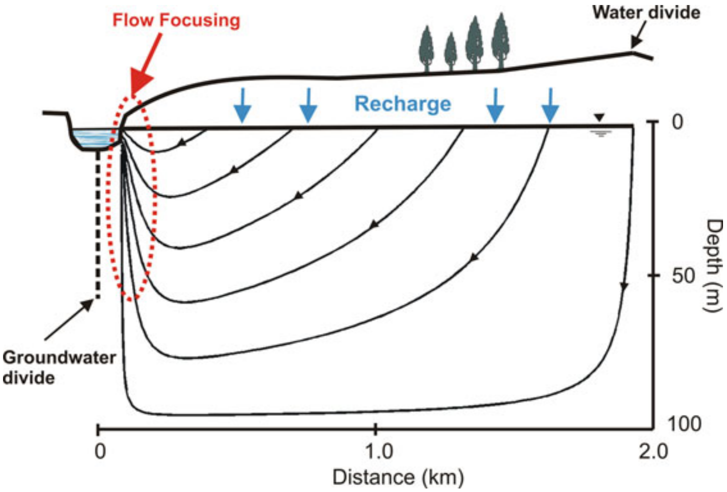


Fig. 19.13 Groundwater flow pattern for a small basin including a river

another example of flow focusing is illustrated. Here a small basin was simulated with the inflow of uniform recharge (365 mm year^{-1}) and the discharge point corresponding to a river located at the groundwater divide. Particle tracking simulations using the code MODPATH (Pollock 1994) clearly show the focusing of path lines close to the river, resulting, also in this case, in a stronger mixing and, therefore, a potentially more efficient reactive zone.

19.5 Case Study: Transport of Ammonium from a Landfill

Examples of scenario specific modeling for the identification of governing processes and importance of parameters have been shown in the previous section. In this section an application of reactive transport modeling for a real case study is presented.The domestic waste landfill site close to the village of "Osterhofen" in south-western Germany has been investigated systematically as monitored Natural Attenuation (MNA) reference site by order of the state EPA (LfU Baden-Württemberg). Like many others, this landfill emits an extended plume of ammonium (NH_4^+) as major contaminant. A detailed characterization of the site and the assessment of monitored Natural Attenuation as land management option can be found in Rügner et al. (2004). The investigation of the site by integral pumping tests is described in Bayer-Raich et al. (2004) and the modeling approach in Maier and Grathwohl (2006).

The former gravel pit was filled with municipal waste between 1969 and 1977. The area is built up by moraine sediments and periglacial deposits from the latest alpine glaciation period. The contaminated gravel aquifer has a thickness between two and five meters. Groundwater flow velocities between two and three meters per day were observed. Annual water table fluctuations between 1.5 and 3.5 m have been measured in the area. In general, the regional aquifer is aerobic with concentrations of oxygen between 6 and 10 mg L^{-1}. Nitrate concentrations in the range of the legal limit (50 mg L^{-1}) were observed in the vicinity of the site, which originate from intensive fertilization in agriculture. Directly down-gradient of the landfill, anaerobic conditions prevail at a width of 150 m with zero concentrations of oxygen and nitrate. Within this area elevated ammonium concentrations were measured between 10 and 40 mg L^{-1}, which are substantially above the legal limit of 0.5 mg L^{-1}.

Multilevel sampling demonstrated that the whole thickness of the aquifer is contaminated by the ammonium plume. Dissolved organic carbon (DOC) was measured less abundant between 1 and 8 mg L^{-1} and contributes to a minor degree (<20%) to oxygen demand compared to NH_4^+. Groundwater samples also indicate that the concentrations of other electron donors such as Fe(II) or Mn(II) are negligible.

The groundwater concentrations and total mass fluxes of ammonium from the landfill were determined by the method of integral pumping tests at two consecutive control planes perpendicular to the flow direction at the site. This method is especially suited to obtain highly reliable information even for large scale investigations (Bayer-Raich et al. 2004). The first control plane was installed directly at the edge of the landfill and the second control plane is approximately 450 m down-gradient in the groundwater flow direction of the landfill. At the second control plane much lower concentrations of NH_4^+ were observed. An overview of the site and two multilevel sampling control planes are given in Fig. 19.14.

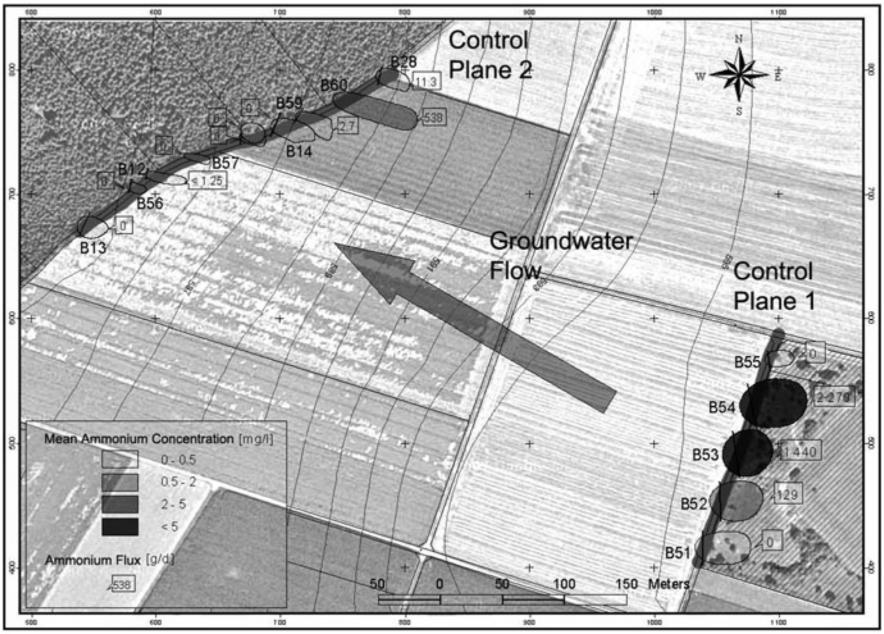


Fig. 19.14 The monitored Natural Attenuation (MNA) reference site "Osterhofen" in SW-Germany. The municipal landfill is emitting an ammonium plume which is monitored at 2 consecutive control planes (modified from Maier and Grathwohl, 2006)

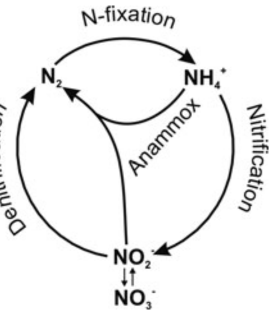


Fig. 19.15 The biogeochemical nitrogen cycle

A first step in the formulation of the conceptual model is the identification of the relevant processes of Natural Attenuation that may lead to the observed elimination of a contaminant. In the case of ammonium, potential biochemical transformations will be part of the nitrogen cycle (Schlesinger 1997), which is depicted as a sketch in Fig. 19.15.

As can be seen in this figure, two possible reaction pathways may lead to NH_4^+ decomposition: nitrification, which requires the presence of oxygen and the anammox reaction as a symproportionation of NH_4^+ and NO_3^-. Furthermore essential for the definition of an appropriate conceptual model is the identification of potential processes that deliver electron acceptors into the zone of contamination. Biotransformation may take place as core controlled processes within the interior of the plume or as fringe controlled processes driven by external electron acceptors. Plume core processes which allow for degradation without external mixing of electron acceptors are strongly dependent on the available electron acceptors, degradation kinetics and stoichiometry. Anaerobic ammonium degradation under reducing conditions using nitrate, manganese or iron oxides as intrinsic electron acceptors (anammox) has been described only recently and may contribute to ammonium degradation to some extent (Buss et al. 2003; Christensen et al. 2001; Jetten 2001; Schink 2002). However, field data indicate that these mechanisms are rather slow and are unlikely to contribute significantly to Natural Attenuation of ammonium. Moreover, manganese oxides, which are described as potential electron acceptors to transfer ammonium to nitrogen in Buss et al. (2003) may be depleted quickly in the aquifer material within the plume area, where also nitrate was shown not to be present. Thus, core processes are unlikely to contribute to Natural Attenuation of NH_4^+; hence they were not considered in the conceptual model for this study. Oxygen for nitrification is not present in the landfill plume. Therefore, delivery of O_2 by vertical mixing from the fringe of the plume is identified as the essential process for NH_4^+ degradation at the site. Quantification of transverse dispersion (Eq. (19.11)) with the transverse dispersivity (α_T) as unknown parameter proves crucial for estimation of Natural Attenuation potential at the site.

Aerobic nitrification of ammonium to nitrate in the presence of oxygen (Stumm and Morgan 1996) could be identified as governing degradation process in the given case. As two different bacterial strains are involved, the reaction proceeds in two steps. The intermediate product nitrite (NO_2^-) is unstable, so that nitrification is commonly combined to the overall reaction (19.23):

$$NH_4^+ + 1.5O_2 = NO_2^- + H_2O + 2H^+ \tag{19.21}$$

$$NO_2^- + 0.5O_2 = NO_3^- \tag{19.22}$$

$$NH_4^+ + 2O_2 = NO_3^- + H_2O + 2H^+ \tag{19.23}$$

The kinetics of biodegradation (i.e. biochemical reaction rates of two components A and B) are commonly expressed according to a multiplicative-Monod formulation. Additionally, threshold terms preventing the reaction at very low concentrations, were included in the model formulation:

$$R = -k_{\max} \cdot \frac{C_A}{C_A + K_{1/2}^A} \cdot \frac{C_B}{C_B + K_{1/2}^B} \cdot \left(\frac{C_B}{C_B + K_{thr}^B}\right)^2 \cdot \left(\frac{C_A}{C_A + K_{thr}^A}\right)^2 \tag{19.24}$$

where A represents ammonium and B oxygen. R [M $L^{-3}T^{-1}$] is the reaction rate at a certain location, $k_{\max}$ [M $L^{-3}T^{-1}$] is the Monod maximum utilization rate of the reaction, $K_{1/2}$ [M L^{-3}] is the Monod half saturation constant of the reaction for each contaminant. In the given case additional terms are introduced that specify required minimum concentrations of substrate and electron acceptor using K_{thr} [M L^{-3}] as the threshold concentration of the specific contaminant that is just sufficient to maintain the reaction.

A frequent prerequisite for acceptance of Natural Attenuation at a site is the proof that the contaminant plume is not advancing anymore. Therefore, evidence for steady state conditions is needed, which is typically given by the observation that the plume is significantly shorter than the travel distance of the groundwater taking into account the age of the source, flow velocity and possible retardation. In order to assess if an observed ammonium plume can be assumed to be at steady state, an estimation of retardation factors may thus be very important. It should be noted, however, that retardation does not affect the length of a steady state plume, but only defines how much time is needed until steady state conditions are achieved.

The major process of retardation of ammonium as an inorganic cation is generally assumed to be ion exchange at negatively charged external and internal surfaces of clay minerals as well as organic matter in the soil and aquifer material. Ion exchange is competitive because exchange sites are limited. Ionic species of highest concentration and activity within the aqueous solution will predominate the exchanger composition, which has to be taken into account when estimating retardation factors of specific species. Potassium (K^+) may easily replace NH_4^+ due to the similar ion radius, however, in most environments it is only present in small concentrations. In the presence of limestone rocks calcium (Ca^{2+}) is the major cation in groundwater aquifers and will therefore dominate in the competition for cation exchange sites. The reaction equation can be formulated as:

$$NH_4^+ + 1/2Ca - X_2 = NH_4 - X + 1/2Ca^{2+} \quad (19.25)$$

According to the law of mass action and the Gaines-Thomas-formulation of ion exchange we obtain as equilibrium constant $K_{NH4\backslash Ca}$:

$$K_{NH4\backslash Ca} = \frac{[NH_4 - X] \cdot [Ca^{2+}]^{0.5}}{[Ca - X_2]^{0.5} \cdot [NH_4^+]} = \frac{\beta_{NH4} \cdot [Ca^{2+}]^{0.5}}{\beta_{Ca}^{0.5} \cdot [NH_4^+]} \quad (19.26)$$

where the values in squared brackets relate to the molar concentration: $[NH_4 - X]$, $[Ca - X_2]$, $[Ca^{2+}]$, and $[NH^{4+}]$ are the ammonium concentration on the exchange sites, the calcium concentration of the exchange sites, and the aqueous concentrations of calcium and ammonium, respectively. β [–] is the fraction of a specific component on the total exchange capacity. The exchange coefficient K_d between aqueous and solid phase and the retardation factor R are (see Appelo and Postma 2005):

$$K_d = \frac{CEC/100}{m} \cdot \frac{\beta_{NH4}}{[NH_4^+]};\ R = 1 + K_d \frac{\rho}{n} \quad (19.27)$$

where *CEC* is the total cation exchange capacity [meq/100 g sediment] and *m* the charge of the ion (for ammonium = 1). ρ denotes soil bulk density [g mL^{-1}], and *n* the porosity [–].

Table 19.2 gives an overview of retardation factors to be expected for waters with different exchange coefficients (*CEC*). Retardation factors appear to depend only weakly on NH_4^+ concentration.

In the literature, retardation factors for ammonium are reported to vary in a range of 1.5–13 for fine grained sediments (e.g. Christensen et al. 2001; Erskine 2000; Haerens et al. 2002). For gravel aquifers, in contrast, cation exchange capacities (CECs) were found to be rather low,i.e.between 0.8 and 1.2 meq/100 g in carbonatic river Neckar and Rhein gravel (Danzer 1999). These values will lead to rather low retardation factors. Thus it can be assumed that the 25 years old plume at the field site with a flow velocity between 2 and 3 m day^{-1} is at steady state and sorption can be neglected in the conceptual model.

The conceptual model is transformed into a numerical model by spatial and temporal discretization using appropriate dimensionality in an appropriate numerical

Table 19.2 Examples of retardation factors for ammonium in groundwater in contact with different sedimentary rocks due to cation exchange competition with calcium

CEC [meq/100 g]	Sand with Ca^{2+} = 10 mg/L R:	Limestone with Ca^{2+} = 150 mg/L R:	Gypsum with Ca^{2+} = 450 mg/L R:
1	5–6	2.2–2.3	1.7
2	9–11	3.4–3.6	2.4–2.5
5	21–26	7–7.4	4.5–4.7

code. Whereas a full three-dimensional model would generally represent all possible geometric boundaries for the involved processes in the most realistic way, it is useful to reduce the dimensionality of the model as far as possible to reduce computational demand. An approach to reduce such models from 2 into 1 dimension by elimination of the flow distance by travel time was outlined in Maier and Grathwohl (2006). Further it is recommendable to use two or more different codes with different numerical approaches and underlying computational methods to improve confidence in the results and to prove their consistency.

Numerical tools have the advantage of their applicability to a wide range of scenarios and can easily be adapted to field conditions. Values of dispersivity that are significantly smaller than grid spacing, however, are a challenge for the accuracy of numerical models. Coarse grids bear the risk of artificial mixing which can lead to overprediction of mixing and Natural Attenuation rates. Several strategies were applied to overcome this potential source of inaccuracy. Different models using different numerical approaches were compared and yielded consistent results. In a sensitivity analysis, the effect of reducing α_T was investigated and grid spacing was carefully refined until no more influence of the choice of grid spacing could be observed.

The width of the contaminant plume of 300 m is considerably larger than its thickness, so lateral transport of oxygen into the plume can be assumed of minor influence. Therefore two-dimensional models in a vertical (xz)-plane were discretized using BIONAPL, an operator-splitting finite element code (Molson et al. 2002) and MIN3P, a global implicit finite volume code (Mayer et al. 2002). The site is assumed to consist of an aerobic aquifer that is contaminated over its whole thickness. An average aquifer thickness of 4.5 m, groundwater flow velocity of 2.7 m/day, porosity of 35%, an average influx concentration of NH_4^+ of 15 mg L^{-1}, and an oxygen concentration at the water table of 8 mg L^{-1} were simulated. Grid spacing was 1 m/0.1 m in x- and z-direction, respectively, resulting in a model domain of 1,000 by 450 elements.

The parameter for the quantification of the amount of mixing within the model domain of the aquifer is transverse dispersivity α_T, which is not known initially. A very important aspect is to extract an aquifer scale value of dispersivity under field conditions. The aim was to reproduce the depth-averaged concentration of NH_4^+ of 1.5 mg L^{-1} encountered at the control plane 450 m down-gradient of the landfill by fitting α_T by repeated model runs until satisfactory agreement was achieved. In the two-dimensional model, the best fit was achieved with transverse vertical dispersivity α_T of 3.2 cm. A length of the plume of 570 m was obtained which agrees with the field observations. Figure 19.16 shows the concentrations of ammonium and oxygen at the aquifer bottom for 250, 350, 500 and 600 days of simulation.

Concentrations at different times show that steady state was already established after about 500 days for a groundwater velocity (v) of 2.7 m day^{-1}. It should be noted that the fitted $\alpha_T = 3.2$ cm is a bulk parameter that accounts for the whole aquifer in the area of investigation, i.e. including possible effects of heterogeneity that cannot be resolved in detail in a large scale investigation.

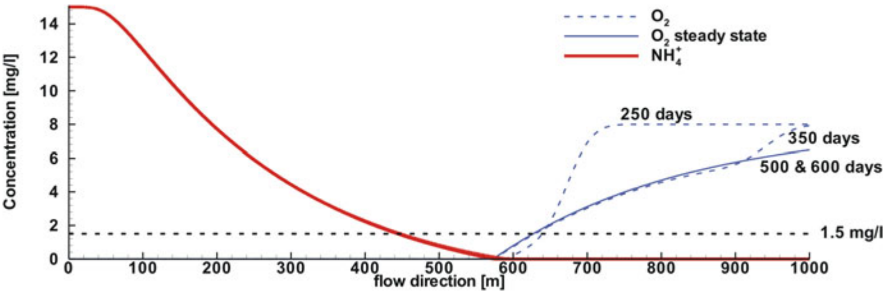


Fig. 19.16 Best fitting results of 2D reactive transport modeling for the Osterhofen site compared to the concentration of 1.5 mg/L measured at 450 m distance

19.6 Summary and Conclusions

This chapter presents the basic theory on contaminant transport, including a description of the principal physical and biogeochemical processes that determine the formation and migration of contaminant plumes in groundwater. Moreover, an introduction to contaminant transport models and their application is given in order to familiarize the reader with the fundamental tools for the quantitative description of the fate and transport of contaminants in aquifer systems. The discussion is complemented with two practical examples concerning the reactive transport of toluene from a LNAPL source and the transport and degradation of ammonium in a landfill leachate plume.

Despite considerable advances in the last three decades in both process understanding and numerical modeling techniques (e.g. improved numerical algorithms and availability of larger computational resources), additional effort is needed in order to fill still existing knowledge gaps and to make progress in our capability to understand and rigorously describe the complex natural processes that determine contaminant transport in natural formations. In the opinion of the authors, active and promising research fields are the investigation of diffusive/dispersive mixing processes in groundwater and their impact on mixing-controlled reactions in homogeneous and heterogeneous formations, together with an improved knowledge and capability of quantitatively describe the impact of the activity of microbial communities on contaminant transport and degradation.

References

Alexander M (1998) Biodegradation and bioremediation. Academic, London

Allen-King R, Grathwohl P, Ball WP (2002) New modeling paradigms for the sorption of hydrophobic organic chemicals to heterogeneous carbonaceous matter in soils, sediments, and rocks. Adv Water Resour 25:985–1016 (also published in Miller et al. 2003 "25 Years of Advances in Water Resources")

Anneser B, Einsiedl F, Meckenstock RU, Richters L, Wisotzky F, Griebler C (2008) High-resolution monitoring of biogeochemical gradients in a tar oil-contaminated aquifer. Appl Geochem 23:1715–1730

Appelo CAJ, Postma D (2005) Geochemistry, groundwater and pollution, 2nd edn. A.A. Balkema, Rotterdam

Baedecker MJ, Back W (1979) Hydrogeological processes and chemical reactions at a landfill. Ground Water 17:429–437

Barry DA, Prommer H, Miller CT, Engesgaard P, Brun A, Zheng C (2002) Modelling the fate of oxidisable organic contaminants in groundwater. Adv Water Resour 25:945–983

Bauer RD, Maloszewski P, Zhang Y, Meckenstock RU, Griebler C (2008) Mixing-controlled aerobic and anaerobic biodegradation of toluene in porous media – results from two-dimensional laboratory experiments. J Contamin Hydrol 96:150–168

Bauer RD, Rolle M, Eberhardt C, Grathwohl P, Bauer S, Kolditz O, Meckenstock RU, Griebler C (2009a) Enhanced biodegradation by hydraulic heterogeneities in petroleum hydrocarbon plumes. J Contamin Hydrol 105:56–68

Bauer RD, Rolle M, Kürzinger P, Grathwohl P, Meckenstock RU, Griebler C (2009b) Two-dimensional flow-through microcosms – versatile test systems to study biodegradation processes in porous aquifers. J Hydrol 369:284–295

Bayer-Raich M, Jarsjo J, Liedl R, Ptak T, Teutsch G (2004) Average contaminant concentration and mass flow in aquifers from time-dependent pumping well data: analytical framework. Water Resour Res 40:W08303. doi:10.1029/2004WR003095

Bear J (1972) Dynamics of fluids in porous media. Dover Publ Inc, New York

Bear J (1979) Hydraulics of groundwater. McGraw-Hill, New York

Bear J, Beljin MS, Ross RR (1992) Fundamentals of ground-water modeling. US EPA/540/S-92/005

Brun A, Engesgaard P, Christensen TH, Rosbjerg D (2002) Modelling of transport and biogeochemical processes in pollution plumes: Vejen Landfill, Denmark. J Hydrol 256:228–247

Buss SR, Herbert W, Morgan P, Thornton SF (2003) Review of ammonium attenuation in soil and groundwater. Environmental Agency, Bristol, UK, p 68

Chapelle FH (2001) Ground-water microbiology and geochemistry. Wiley, New York

Chapelle FH, McMahon PB, Dubrovsky NM, Fuji RF, Oaksford ET, Vroblesky DA (1995) Deducing the distribution of terminal electron-accepting processes in hydrologically diverse groundwater systems. Water Resour Res 31:359–371

Chiogna, G, Eberhardt, C, Grathwohl, P, Cirpka, OA, Rolle, M (2010) Evidence of compound dependent hydrodynamic and mechanical transverse dispersion by multi-tracer laboratory experiments. Environ Sci Technol 44:688–693

Christensen TH, Bjerg PL, Banwart SA, Jakobsen R, Heron G, Albrechtsen HJ (2000) Characterization of redox conditions in groundwater contaminant plumes. J Contamin Hydrol 45:165–241

Christensen TH, Kjeldsen P, Bjerg PL, Jensen DL, Christensen JB, Baun A, Albrechtsen HJ, Heron C (2001) Biogeochemistry of landfill leachate plumes. Appl Geochem 16(7–8):659–718

Cirpka OA (2005) Effects of sorption on transverse mixing under transient flow. J Contamin Hydrol 78:207–229

Cirpka OA, Frind EO, Helmig R (1999) Numerical simulation of biodegradation controlled by transverse mixing. J Contamin Hydrol 40:159–182

Cirpka OA, Schwede RL, Luo J, Dentz M (2008) Concentration statistics for mixing controlled reactive transport in random heterogeneous media. J Contamin Hydrol 98:61–74

Clement TP (1997) RT3D – a modular computer code for simulating reactive multispecies transport in 3-dimensional groundwater aquifers. Battelle Pacific Northwest National Laboratory, PNNL-SA-28967, Richland, USA

Cookson JT (1995) Bioremediation Engineering; design and applications. McGraw-Hill, New York

Dagan G (1989) Flow and transport in porous formations. Springer, Berlin

Danzer J (1999) Surfactant transport and coupled transport of polycyclic aromatic hydrocarbons (PAHs) and surfactants in natural aquifer material laboratory experiments. Tübinger Geowissenschaftliche Arbeiten (TGA). Tübingen, Attempto, 75p. http://w210.ub.uni-tuebingen.de/dbt/volltexte/2005/2019/pdf/49.pdf

Eberhardt C, Grathwohl P (2002) Time scales of organic contaminant dissolution from complex source zones: coal tar pools vs blobs. J Contamin Hydrol 59:45–66

Erskine AD (2000) Transport of ammonium in aquifers: retardation and degradation. Q J Eng Geol Hydrogeol 33:161–170

Fetter CW (1993) Contaminant hydrology. Prentice-Hall, New York

Gelhar LW (1993) Stochastic Subsurface Hydrology. Prentice Hall, Englewood Cliffs, NJ

Grathwohl P (1998) Diffusion in natural porous media: contaminant transport, sorption/desorption and dissolution kinetics. Kluwer Academic, New York

Haerens B, Dassargues A, Prommer H, Lerner D (2002) Multicomponent reactive transport modelling of ammonium contamination at a former coal carbonisation plant. Aardkundige Mededelingen 12:211–214

Ham PAS, Schotting RJ, Prommer H, Davis GB (2004) Effects of hydrodynamic dispersion on plume lengths for instantaneous bimolecular reactions. Adv Water Resour 27:803–813

Hill MC, Tiedeman CR (2007) Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty. Wiley, Hoboken, NJ

Hunter KS, Wang Y, Van Cappellen P (1998) Kinetic modelling of microbially-driven redox chemistry of subsurface environments: coupling transport, microbial metabolism and geochemistry. J Hydrol 209:53–80

Imhoff PT, Jaffe PR, Pinder GF (1993) An experimental study of complete dissolution of a nonaqueous phase liquid in saturated porous media. Water Resour Res 30:307–320

Jetten MSM (2001) New pathways for ammonia conversion in soil and aquatic systems. Plant Soil 230:9–19

Klenk ID, Grathwohl P (2002) Transverse vertical dispersion in groundwater and the capillary fringe. J Contamin Hydrol 58(1–2):111–128

Lerner DN, Thornton SF, Spence MJ, Banwart SA, Bottrell SH, Higgo J. J, Mallison HEH, Pickup RW, Williams, GM (2000) Ineffective natural attenuation of degradable organic compounds in groundwater. Ground Water 38:922–928

Liedl R, Valocchi AJ, Dietrich P, Grathwohl P (2005) Finiteness of steady-state plumes. Water Resour Res 41:1–8

Maier U, Grathwohl P (2006) Numerical experiments and field results on the size of steady state plumes. J Contamin Hydrol 85:33–52

Mayer KU, Frind EO, Blowes DW (2002) Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. Water Resour Res 38:1174–1195

McDonald MG, Harbaugh AW (1988) A modular three-dimensional finite-difference ground-water flow model. US Geological Survey Techniques of Water Resources Investigations, Book 6

Miller CT, Poirier-McNeill MM, Mayer AS (1990) Dissolution of trapped nonaqueous phase liquids: mass transfer characteristics. Water Resour Res 26:2783–2796

Molson JW, Frind EO, Van Stempvoort DR, Lesage S (2002) Humic acid enhanced remediation of an emplaced diesel source in groundwater. 2. Numerical model development and application. J Contamin Hydrol 54:277–305

Olsson AH, Grathwohl P (2007) Transverse dispersion of non-reactive tracers in porous media: a new nonlinear relationship to predict dispersion coefficients. J Contamin Hydrol 92: 149–161

Pankow JF, Cherry JA (1996) Dense chlorinated solvents and other DNAPLs in groundwater. Waterloo Press, Portland Oregon

Parkhurst DL, Appelo CAJ (1999) User's guide to PHREEQC – a computer program for speciation, reaction-path, 1D-transport, and inverse geochemical calculations. Technical Report 99–4259, US Geol. Survey Water-Resources Investigations Report

Pollock DW (1994) User's guide for modpath/modpath-plot, version 3: a particle tracking post-processing package for modflow, the U.S. geological survey finite-difference ground-water flow model. Open-file Reports 94–464, US Geol Survey, Reston, VA

Prommer H, Barry DA (2005) Modeling bioremediation of contaminated groundwater. In: Bioremediation: applied microbial solutions for real-world environmental cleanup. American Society for Microbiology, Washington, DC, pp 108–138

Prommer H, Barry DA, Davis GB (2002) Modelling of physical and reactive processes during biodegradation of a hydrocarbon plume under transient groundwater flow conditions. J Contamin Hydrol 59:113–131

Prommer H, Barry DA, Zheng C (2003) PHT3D – a MODFLOW/MT3DMS based reactive multicomponent transport model. Ground Water 42:247–257

Rolle M, Bauer R, Griebler C, Meckenstock R, Bauer S, Kolditz O, Grathwohl P (2008b) Aerobic degradation of a toluene plume in homogeneous and heterogeneous porous media. IAHS Red Book Series, 324:356–363

Rolle M, Clement TP, Sethi R, Di Molfetta A (2008a) A kinetic approach for simulating redox-controlled fringe and core biodegradation processes in groundwater: model development and application to a landfill site in Piedmont, Italy. Hydrol Process 22:4905–4921

Rolle M, Eberhardt C, Chiogna G, Grathwohl P (2009) Enhancement of dilution and transverse reactive mixing in porous media: experiments and model-based interpretation. J Contamin Hydrol 110:130–142

Rolle M, Eberhardt C, Grathwohl P (2008c) Influence of flow focusing on transverse mixing in heterogeneous porous media: laboratory experiments and model-based interpretation. IAHS Red Book Series 324:348–355

Rolle M, Piepenbrink M, Prommer H, Ptak T, Grathwohl P (2005) Contaminant transport scenario modelling as a tool for a critical evaluation of natural attenuation. Proceedings of the aquifer vulnerability and risk 2nd internationlal workshop, Reggia di Colorno (PR), 21–23 September

Rügner H, Holder T, Maier U, Bayer-Raich M, Grathwohl P, Teutsch G (2004) Natural attenuation untersuchungen "ehemalige abfalldeponie osterhofen". Grundwasser 9:98–108

Schink B (2002) Anaerobic digestion: concepts, limits and perspectives. Water Sci Technol 45:1–8

Schlesinger WH (1997) Biogeochemistry: an analysis of global change. Academic, New York

Schwarzenbach RP, Gschwend PM, Imboden DM (1993) Environmental organic chemistry. Wiley, New York

Steefel CI, MacQuarrie KTB (1996) Approaches to modeling of reactive transport in porous media. In: Lichtner PC, Steefel CI, Oelkers EH (eds) Reactive transport in porous media, volume 34 of reviews in mineralogy. Mineralogical Society of America, Washington, DC

Stumm W, Morgan JJ (1996) Aquatic chemistry, 3rd edn. Wiley Interscience, New York

Werth CJ, Cirpka OA, Grathwohl P (2006) Enhanced mixing and reaction through flow focusing in heterogeneous porous media. Water Resour Res 42:W12414. doi:10.1029/2005WR00451

Wiedemeier TH, Rifai HS, Newell C, Wilson JT (1999) Natural attenuation of fuels and chlorinated solvents in the subsurface. Wiley, New York

Yeh GT, Ward DS (1980) A finite-element model of water flow through saturated-unsaturated media. ORNL-5567, Oak Ridge National Laboratories, Oak Ridge, TN

Zheng C, Bennet GD (2002) Applied contaminant transport modeling, 2nd edn. Wiley, New York

Zheng C, Wang PP (1999) MT3DMS: a modular three-dimensional multispecies model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems. Documentation and User's Guide. Contract Report SERDP-99–1, US Army Engineer Research and Development Center, Vicksburg, MS

Part VI
Risk Management

Chapter 20
Sustainability and Remediation

R. Paul Bardos, Laurent M.M. Bakker, Hans L.A. Slenders, and C. Paul Nathanail

Abstract *Sustainable remediation* has come to exist as a popular term used to describe contaminated site management that is demonstrably sustainable, i.e. where some form of sustainability appraisal has been used in decision making to identify the "most sustainable" approach for any particular management intervention required. The "most sustainable" approach is one that, in the view of the stakeholders involved in making or considering management decisions, has the optimal balance of effects and benefits across the three elements of sustainability: environment, economy and society. This chapter describes how the Brundtland Report concept of sustainable development can be linked with contaminated site remediation practice, both how sustainability can be assessed and used as a tool in decision making; and also how sustainability thinking is creating new contaminated site remediation approaches, for example, marrying concepts of Risk Management and renewable energy production. The chapter discusses the individual "indicators" or metrics that contribute to an understanding of sustainability, the tools available for combining these into a sustainability appraisal, and the types of boundary conditions that need to be considered in setting the scope of sustainability appraisal. The chapter also discusses the linkage of "sustainability" with "Risk Management", the importance of engaging stakeholders in sustainability appraisal, and an emerging set of international initiatives in the field. Finally it presents a series of sustainable remediation case studies (technologies and decision making tools) and a view of the possible future for "sustainable remediation".

R.P. Bardos (✉)
Environmental Technology Ltd, Reading, UK
e-mail: paul@r3environmental.co.uk

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_20,

Contents

20.1 Introduction

Sustainable development as a concept was defined in the 1987 “Brundtland Report” by the World Commission on Environment and Development as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (Brundtland 1987). This report was an important step on a continuing international debate on sustainable development summarised in Fig. 20.1.

It has long been assumed that contaminated land management was by its nature intrinsically sustainable because, for example, it controlled risks from contaminants and facilitated the re-use of Brownfield land so reducing greenfield development pressure. However, over the past decade it has increasingly been realised that this simple assumption may not always be true. For example, increasing concerns about

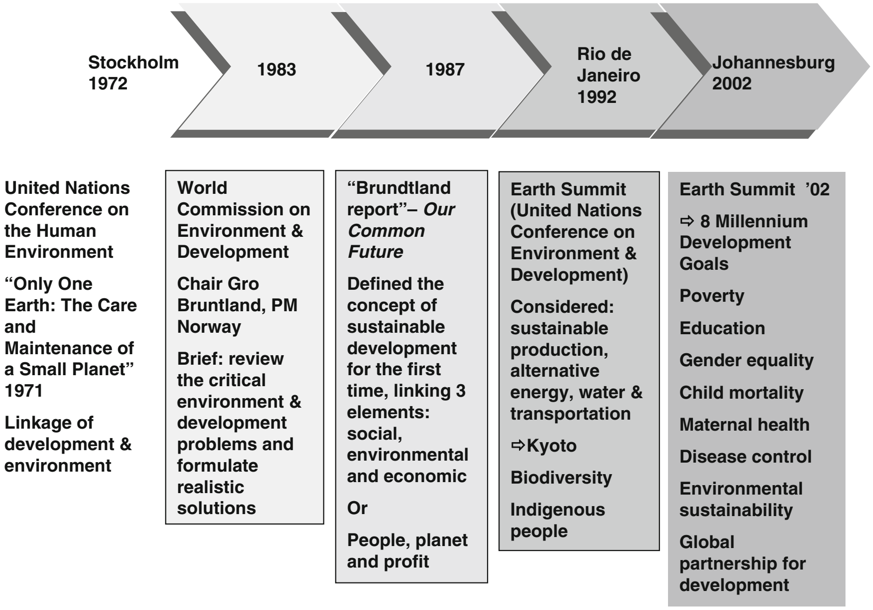


Fig. 20.1 The international sustainable development debate

the use of fossil fuels (and CO_2 emissions) have led to questioning about whether it is truly sustainable to apply energy intensive remediation processes to relatively low levels of contamination. Remediation that does not deliver tangible benefits in the form of reduced risk cannot be said to be a wise use of resources and therefore cannot be seen to be sustainable. Another dimension to this debate has arisen in response to the emergence of interest in the use of marginal land for alternative energy production. For example, in 2008 the US EPA (United States Environmental Protection Agency) launched a national map linking renewable energy opportunities to contaminated sites and mining sites.[1] This chapter explores, conceptually, what might be meant by the term "sustainable remediation"; links this concept to frameworks for contaminated site management and explores how sustainability can be considered in contaminated land or site management decision making and practice, concluding with a selection of case studies.

20.2 Concepts

20.2.1 Sustainability

20.2.1.1 Sustainability Appraisal in Overview

Sustainability is perhaps best described as a measure of how well a particular endeavour is able to meet the goals of sustainable development, for example as

---

[1] http://www.epa.gov/renewableenergyland/

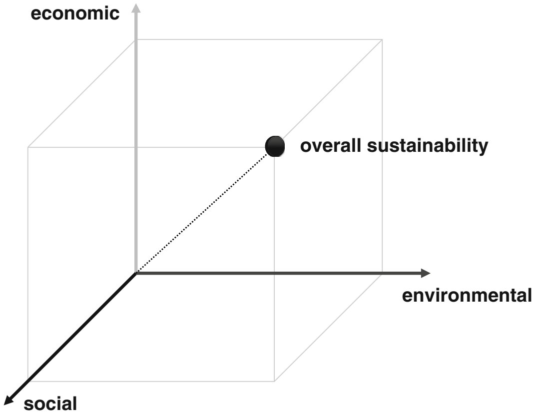


Fig. 20.2 The three elements of sustainable development

defined in the 1987 Brundtland Report. There are three elements to sustainable development, and hence sustainability, as shown in Fig. 20.2: environment, economy and society.

There is no "standard" technique for sustainability appraisal (for example like ISO 14040 for Life Cycle Assessment). Generally, sustainability appraisal tends to be based on assessments of indicators. Indicators are metrics or assessments of individual factors that contribute to an overall understanding of sustainability, for example: direct costs, greenhouse gas emissions etc (CL:AIRE 2009). Sustainability appraisal techniques employ some means of aggregating individual assessments of indicators to provide an overall understanding of "sustainability". Qualitative and quantitative approaches may be used in sustainability appraisal. The more widely used techniques are summarised in Table 20.1, which also highlights which techniques have been applied in contaminated site management (CSM), and to what extent each technique assesses each of the three elements of sustainability. Sustainability appraisal methods are described in more detail in Section 20.3.4. In general, quantitative approaches are limited to particular aspects of sustainability, but may be useful for evidence gathering as part of an overall appraisal. It is also worth pointing out that environmental Risk Assessment is a technique which has a place within sustainability appraisal, but for contaminated site management it is also the dominant basis for decision making as discussed in Section 20.2.2.

20.2.1.2 Using Indicators in Sustainability Appraisal

Sustainability appraisal is often linked to particular policy perceptions and objectives. This means that there are two broad approaches to identifying individual sustainability indicators:

- policy orientated indicators that are linked to specific policy goals, often with some threshold or target for "acceptability" included, and

Table 20.1 Decision support techniques with relevance to sustainability appraisal

	Env	Ec	So	Type	CSM
Scoring/ranking systems (including multi-criteria analysis)	Narrow to Wide	Narrow to Wide	Narrow to Wide	Qual	yes
Best Available Technique (BAT)	Narrow to Wide	Narrow		Qual	yes
Carbon footprint ("area")	Narrow			Quan	yes
Carbon balance (flows)	Narrow			Quan	
Cost benefit analysis	Narrow to Wide	Narrow to Wide	Narrow to Wide	Quan	yes
Cost effectiveness analysis	Narrow to Wide	Narrow to Wide	Narrow to Wide	Qual	yes
Eco-efficiency	Narrow			Quan	?
Ecological footprint	Narrow			Quan	
Energy / intensity efficiency	Narrow			Quan	yes
Environmental Risk Assessment	Narrow to Wide			Quan	yes
Human Health Risk Assessment			Narrow		yes
Environmental impact assessment / Strategic environmental assessment	Narrow to Wide			Qual	yes
Financial Risk Assessment		Narrow		Quan	yes
Industrial ecology	Narrow to Wide	Narrow to Wide		Quan	
Life Cycle Assessment (based)	Narrow to Wide			Quan	yes
Quality of life assessment	Wide	Wide	Wide	Qual	

Indicating coverage of the environmental (En), economic (Ec) and social (So) elements of sustainable development; whether techniques are quantitative or qualitative; and whether contaminated site management – CSM -applications are known to exist at present – see Section 20.3.4.
Coverage: The table describes each technique in terms of its *typical* coverage of particular aspects of sustainability, in terms of the categories set out in Table 20.2 or 20.4. For example, a carbon footprint appraisal focuses on a "narrow" segment of environmental sustainability issues (ignoring for example soil functionality, Biodiversity and landscape impacts), whereas all of these aspects could be considered by a Cost benefit analysis, providing it was suitably specified. Where no entry is made the technique has no coverage

- indicators (metrics?) that are orientated towards consistent reporting of sustainability effects, independent of particular regional, national or international policy goals.

Policy Orientated Indicators

Target or *policy orientated indicators* relate directly to explicit goals. They tend to be identified on the basis of an idea that is considered by stakeholders as being sustainable. An indicator is selected that is representative of this idea, and a target is set to indicate a policy threshold indicating a satisfactory level of performance. An example of such an idea from England is: *avoidance of greenfield development*; indicator: *new houses built on previously developed land*; target: *2008 target of building over 60% of additional housing on previously developed land and conversions of existing buildings (*Department for the Environment Transport and the Regions 1998). The presumption is that the more targets that are met the better sustainability overall for a particular policy sector, or across (say) a national set of sustainability indicators such as the *England Sustainable Development Policy Framework Indicators* (Department for Environment Food and Rural Affairs 2005a, 2005b).

The advantages of target or policy orientated indicators are that there is a specific vision of "sustainability" and that there is political "ownership", i.e., direct linkage with policy goals and democratically accountable bodies. These advantages appeal particularly to Public Sector bodies such as planners, regulators, local authorities and public sector funders. A disadvantage of using these indicators is that they are less likely to provide a "holistic" understanding of sustainability or a consistent basis for reporting, for example to an international audience. For example, the indicators used may have a narrow scope (being focused on particular policy interests and goals). Sustainability appraisals based on these indicator sets tend to be geographically and time limited and vulnerable to political change. They can also be seen as subjective (policy value judgements) and over-simplistic. For example, the previously mentioned English Brownfield target for residential housing used an indicator definition that included gardens as previously developed land, which has led to what many see as inappropriate development of suburban gardens in England, including construction of blocks of flats or high-density housing in areas of otherwise low-density housing (Bennett 2006).

Indicators for Cross Sectoral Reporting

Corporate reporting indicators are often identified with the aim of finding a consistent approach to reporting on sustainability measures, for example across sectors or industries. An example of these is the *Global Reporting Initiative's* (GRI) *indicator*

Table 20.2 2008 listing of GRI environmental, economic and social sustainability criteria

Environmental	Economic	Social
Materials Energy Water Biodiversity Emissions, effluent and waste Products and services Compliance Transport Overall	Economic performance Market presence Indirect economic impacts	Labour practices Human rights Society Product responsibility. *Each of these are further divided into subsidiary aspects and then individual indicators.*

sets.[2] GRI (2006) has pioneered the development of an international sustainability reporting framework.[3] Table 20.2 lists Environmental, economic and social sustainability aspects proposed by GRI, which illustrates the wide coverage of the GRI indicator set. Another example of an international indicator set has been published on line by the UN Department of Economic and Social Affairs Division of Sustainable Development.[4] Individual effects may be impacts or benefits which are important in the context of sustainable development. A broad range of effects establishes the indicator set of interest.

Advantages of using a reporting orientated approach to selecting indicators for sustainability appraisal include the explicit intention to provide a holistic vision of "sustainability", and a greater feasibility for international use, harmonisation and protocols, consequently perhaps greater objectivity. These indicator sets may be less geographically limited, at least not for geopolitical reasons, and perhaps have a reduced vulnerability to political changes. A disadvantage is that they may lack the strengths of policy or target orientated indicators, i.e. that they may not necessarily be linked to direct policy drivers / measurables or public sector organisational goals, and may therefore be considered as more esoteric by some users. An example of an industry specific indicator set that has already been developed is that used by the UK water industry for corporate reporting across all functions in their sector (Water UK 2008).

Clearly, the themes covered by particular "policy-orientated" or "reporting orientated" indicator sets may be similar. The main differences in approach are in their functionality and intended audiences. Perhaps, an ideal sustainability appraisal approach is one that is able to separately consider several indicator sets, and then compare and contrast findings for different sets. In both cases the original purpose of the indicator set may not have been for option appraisal, but rather for reporting progress towards policy targets or for corporate reporting, with indicators being considered individually, rather than being aggregated and compared against different management options. Nonetheless, they may be useful tools to provide suggestions for consideration in option appraisal by decision makers.

Indicator sets may be proposed for option appraisal on a project specific basis, for example as suggested by the EC Framework Programme 5 (FP5) RESCUE project Sustainability Assessment Tool (SAT) for Brownfield regeneration (Edwards et al. 2005). It explicitly avoids the term "indicator", but uses the words "aspects" and "priorities" to mean much the same thing and suggests how these might be determined by questionnaires and workshop activities. It is likely, given the prevalence of sustainable development policies in planning, that such a project specific indicator set would be based on an existing related indicator set in policy.

---

[2] www.globalreporting.org

[3] http://www.globalreporting.org

[4] http://www.un.org/esa/sustdev/ see indicator set at: http://www.un.org/esa/sustdev/natlinfo/indicators/isdms2001/table_4.htm

20.2.1.3 The Driver, Pressure, State, Impact and Response (DPSIR) Framework

From an environmental policy setting perspective the *Driver, Pressure, State, Impact and Response* (DPSIR) framework is important. This framework, shown in Fig. 20.3, is a system for analysing. The *Drivers* forces responsible for change such as industrialisation, growing population. These cause *Pressures* such as increasing soil sealing, which affect the *State* of the environment, for example reduced ability to store rainfall. The result of this are *Impacts* such as greater risk of flooding, to which Society may need to make a *Response*. Indicators can describe the state of sustainability (e.g. the health of the population); pressures on sustainability (e.g. emissions of contaminants); responses to problems (e.g. % of remediation projects involving soil re-use); and contexts that decision-makers have little control over (e.g. population structure). The DPSIR concept can also be extended to economic and social indicators (Therivel 2004).

Therivel (2004) points out that most indicator lists include a combination of state, pressure, response and context indicators with the aim of producing a full "picture" of sustainability as possible. Indicator sets are not often explicit about what DPSIR segments they relate to, and as many are mixed there is a possibility of double counting factors. For example, one indicator might be the number of heavy goods vehicle (HGV) miles travelled for a particular remediation project, whilst another might be concerned with global air quality considering NO_x and SO_x across the project (*State)*. However, HGV movements and HGV exhausts are a source of NO_x and SO_x (*Pressure*) so there appears to be double counting in the indicator set. One way of dealing with double counting is to strip down the number of indicators so that only one type of indicator, e.g. *State* and limiting the assessments made to a number of key criteria, e.g. "air quality" as determined by NO_x and SO_x measurements. This is the type of approach that might be adopted for a Life Cycle Assessment type of appraisal. However, this *reductionist approach* has some major disadvantages as shown below.

HGV miles measures more than simply air quality. It includes a rich mixture of impacts on the environment, but also on society (e.g. road safety, noise, disturbance) and economics (fuel costs, costs related to traffic congestion caused). It has a particular and unique characteristic which may be of interest to decision makers. It is

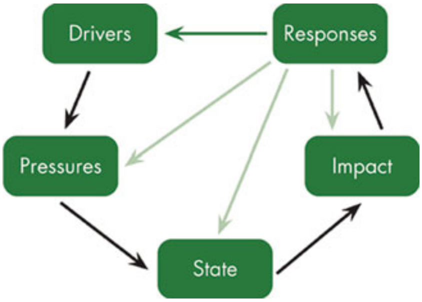


Fig. 20.3 DPSIR Model

also more intellectually accessible to a wider range of stakeholders than a complex set of calculations that convert all factors into a limited set of quantitative metrics.

So there appears to be a dilemma between using decision criteria that are strictly rational, measurable and comparable, and using criteria that are "rich" in their coverage of impacts and may possibly lead to apparent duplication. The way out of this dilemma is that where possible a richer and more accessible set of indicators should be used, but with explanation and transparency, where there may be possible overlapping considerations, or points of cross-reference. For example, one set of indicators might be direct emissions of NOx and SOx of the remediation process, but with a cross reference to state that this assessment specifically excludes impacts associated with emissions from HGV transport, as HGV miles are covered by a separate indicator.

20.2.2 Risk Management

20.2.2.1 Risk Management Principles

Risk Management involves taking practical steps to mitigate identified risks so that they are eliminated or at least reduced to an acceptable level. In land contamination remediation terms it involves demonstrably breaking the source-pathway-receptor contaminant linkage (Nathanail et al. 2007), as shown in Fig. 20.4. Breaking the contaminant can involve removing or otherwise modifying the source, interrupting the pathway or modifying the behaviour of the receptor, or relocating the receptor. Source management can involve contaminant destruction, detoxification, immobilisation, transfer or reduction in concentration. Pathway interruption can involve isolation (e.g. encapsulation) or plume treatment (e.g. monitored Natural Attenuation or permeable reactive barriers). Receptor management can include relocation of sensitive species or in a work place scenario provision of personal protective equipment, restricting access or minimising exposure duration.

Risk Management options for soil and groundwater contamination can be grouped into civil engineering, biological, chemical, physical, solidification/stabilisation, thermal or institutional categories (Nathanail and Bardos 2004). Civil engineering approaches either relocate the contaminant (soil vapour extraction or off site disposal) or physically isolate it from receptors of concern (e.g. vertical hydraulic barriers or capping layers). Biological agents (e.g. microbes or plants)

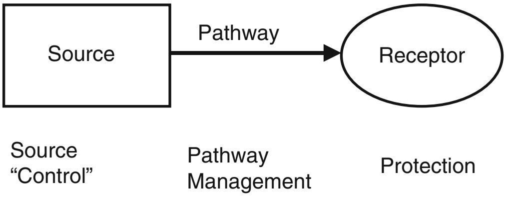


Fig. 20.4 Risk Management interventions by breaking the source-pathway-receptor contaminant linkage

can be applied to destroy contaminants in situ (e.g. bioventing, phytoremediation) or ex situ (e.g. biopiles). Chemical methods destroy (e.g. in situ chemical oxidation), transform or stabilise contaminants. Physical methods exploit differences in soil or contaminant properties such as particle size distributions (soil differences) or density or volatility (contaminant differences) to separate contaminated and uncontaminated fractions and, hence, reducing the volume of material that needs further treatment or disposal. Thermal methods can enhance other technologies (e.g. six phase soil heating can improve the efficiency of soil vapour extraction) or destroy (e.g. incineration), relocate (thermal desorption), or isolate (vitrification) contaminants. Solidification/stabilisation treatments are generally chemical and physical measures which reduce the physical accessibility and/or chemical availability of contaminants. See for more details on in situ remediation technologies Grotenhuis and Rijnaarts (Chapter 21 of this book). For more details on Natural Attenuation see Chapter 22 by Peter et al., this book.

An important tool in planning and verifying Risk Management is the *site conceptual model*. The conceptual model comprises a plan, cross section and topological (network or matrix diagram) graphic with associated, often tabulated, text about the sources, pathways, receptors and associated uncertainties or assumptions. The network diagram provides a simple template on to which components of the Risk Management strategy can be placed to permit a simple check that all the contaminant linkages have been broken. The use of the network diagram in this way can facilitate the design of cost effective, socially acceptable and environmentally benign integrated solutions and also facilitates communication about potential risks between the different stakeholders involved in a particular project.

Typically most remediation work has been initiated for one or more of the following reasons:

- to protect human health, or the environment (including groundwater);
- to enable redevelopment, including restoration of sites for amenity, or bioenergy purposes;
- to limit potential liabilities, for example to improve corporate balance sheets or to facilitate sale or transfer;
- to repair or enhance previous remediation work that has been found to be inadequate.

In all cases the underlying stimulus or intention is to mitigate risks to human health, the environment, groundwater and surface waters and buildings, as generally set out in the prevailing legislative and regulatory framework. Mitigation of these risks is therefore a prerequisite, without which the remediation is essentially functionless, and hence *de facto* unsustainable. Risk Management must be achieved and evidence must be captured and reported that demonstrates that it has been broken to all stakeholders (regulators, future purchaser or tenant, other stakeholders). For example, the UK policy view is that: *a fundamental principle of sustainable development is that the condition of land, its use and its development should be protected from potential hazards. Without appropriate action, the presence of contaminants* *with the potential to cause harm to human health, property and the wider environment may severely limit or altogether preclude development and the beneficial use of land* (Office of the Deputy Prime Minister 2004). Hence, contaminated site management decisions are based on Risk Management, as expressed in the prevailing regulations for the particular region the site is in. In other words, the Risk Management decision sets the *scope* of any remediation work. However, in most circumstances in Europe the Risk Management decision itself is related to the proposed end-use of the site. However, it is increasingly being recognised that the execution of any Risk Management required also needs to be sustainable. Under current circumstances, the "sustainable remediation" debate therefore centres on how to find a reasonable balance between sustainability, Risk Management and land use.

A wider question is whether the environmental quality objectives set by the prevailing regulatory regime or legislative regime are themselves sustainable, but this question is outside the scope of this chapter. For example, some regimes, such as those related to contamination prevention, take a hazard management approach in the case of new contamination. For "recent", "new" or "future" contamination, removal of contamination to background or original levels (rather than a risk based environmental quality criteria) may be required. The Environmental Liability Directive and the Integrated Pollution Prevention and Control Directive is based on the premise that site contamination should be seen as a historic problem and that industrial processes will be managed in ways that do not result in releases to the environment. Where such releases do occur these Directives require their effects to be reversed.

20.2.2.2 Institutional Controls

Institutional controls describe control systems put in place to manage future actions by institutions. These have a bearing on contamination in several ways:

1) Planning policies such as zoning affect the range of potential re-uses for Brownfield sites
2) Controls on future land use (either in planning, or through contractual routes or covenants) that limit the actions or access of a property owner, tenant or other kind of user or person on a particular area of land. These are a key component o the "fit for purpose" concept in risk based land management. Without institutional controls fitness for purpose could not function as a viable approach in the longer term.
3) Requirements for record keeping (for example regarding site condition, past condition and any site investigation, Risk Assessment and Risk Management actions that might have been carried out).

Typically, the concept most commonly referred to is the second one. For example the US EPA defines institutional controls in broad terms as: "legal measures that limit human exposure by restricting activity, use, and access to properties with residual contamination. This site provides a clearinghouse of information on LUCs for the use of all stakeholder groups."[5] However, all three have an important bearing on contaminated land management, and the determination of past and future liabilities.

A common theme for all three types of control is need for the preservation of documentation, and a major concern raised for "fit for purpose" remediation for some jurisdictions is whether public record keeping and contractual record keeping by institutions is adequate. Another issue is related to the preservation of information about sites; as information availability may be a large influence in determining liabilities.

Institutional controls also have a bearing in sustainable development, as already discussed in 2002 in CLARINET's Risk Based Land Management Concept (Vegter et al. 2002). It may be more appropriate to affect a limitation in land use access and activity to allow a lower input approach to remediation in the short term, and accept that a land use change will require further Risk Management. This provides an alternative land use rationale to the use of a multi-functional approach. A partial reason for some sites might be the hope of better and more effective remediation approaches in the future for particularly difficult contamination problems. However, a contingent part of this rationale is the ability to preserve information into the future for long periods, effectively over future generations. In the Netherlands and UK national organisations have been set up for the long term care of Brownfield sites and for the Netherlands for sites with residual contamination after remediation, in return for an initial payment. This is then invested and pays for ongoing site maintenance.[6]

20.2.3 Sustainable Remediation

Sustainable remediation has come to exist as a popular term used to describe contaminated site management that is demonstrably sustainable, i.e. where some form of sustainability appraisal has been used in decision making to identify the "most sustainable" approach for any particular management intervention required. The "most sustainable" approach is one that, in the view of the stakeholders involved in making or considering management decisions, has the optimal balance of effects and benefits for each of the three elements of sustainability: environment, economy and society. A number of international and national initiatives are developing "sustainable remediation" concepts and their provisional definitions are summarised in Table 20.3. These initiatives are described in Section 20.2.4.

Remediation is a process that takes place after a chain of decisions that set its *scope*. Very often the remediation work is part of a larger initiative, for example the redevelopment of a former industrial site, which will include a wide range of other decisions related to feasibility of developing a site and financing the

---

[5] http://www.epa.gov/brownfields/tools/tti_lucs.htm

[6] www.landrestorationtrust.org.uk

Table 20.3 Definitions of "sustainable remediation"

NICOLE[a]	A Sustainable Remediation project is one which the stakeholders agree represents the best solution considering environmental, social and economic factors.
SuRF-UK[b]	The working definition in November 2008: *The practice of demonstrating, in terms of environmental, economic and social indicators, that an acceptable balance exists between the effects of undertaking the remediation activities and the benefits the same activities will deliver.*
SURF-USA[c]	SURF-US' working concept, outlined in the SURF White Paper (SURF 2009) is as follows: • In fulfilling our obligations to remediate sites to be protective of human health and the environment we will embrace sustainable approaches to remediation that provide a net benefit to the environment. • To the extent possible, these approaches will: ○ minimize or eliminate energy consumption or the consumption of other natural resources; ○ reduce or eliminate releases to the environment, especially to the air; ○ harness or mimic a natural process; ○ result in the reuse or recycling of land or otherwise undesirable materials; ○ encourage the use of remediation technologies that permanently destroy contamination.
US Environmental Protection Agency (2008)	"Green remediation" is "the practice of considering all environmental effects of remedy implementation and incorporating options to maximize net environmental benefit of cleanup actions." Green remediation considers a range of impacts: air pollution caused by toxic or priority pollutants such as particulate matter and lead; water cycle imbalance within local and regional hydrologic regimes; soil erosion and nutrient depletion as well as subsurface geochemical changes; ecological diversity and population reductions; and emission of carbon dioxide (CO_2), nitrous oxide (N_2O), methane (CH_4), and other greenhouse gases contributing to climate change.

[a] www.nicole.org
[b] www.claire.co.uk/surfuk
[c] www.sustainableremediation.org

project, and hence their consequent Risk Management requirements. These project based decisions are affected by higher level decisions such as: setting national and regional spatial policies, and indeed Risk Management priorities, and local level decisions such as a municipality's development zoning requirements. Each of these decisions might be subject to its own sustainability appraisal and sustainable development context. Hence, remediation decision making is one segment of a broader framework of decision making as explained in Section 20.2.3.

In other words, a range of wider management decisions also affect the scope of remediation work and its sustainability appraisal. These depend on the site circumstances, the most important of which is likely to be the present use of the site and the development needs that imposes. These decisions determine the scope of possible remediation approaches in two ways:

1. In terms of regulatory and planning controls on environmental risks say to human health, water and the wider environment, i.e. the needs are those that relate to the desired end use of the site.
2. By setting practical boundaries such as the time and space available to carry out remediation, but also limit the range of possible interventions.

Therefore the impact of considering "sustainable remediation" depends on where in the decision making process it takes place. Sustainable remediation may be considered purely as an aspect of selecting the optimum remediation strategy for a project where all land use and site development decisions have already been made. This may often be the scenario that is faced by contaminated site management service providers. In this situation, the impact of adopting a sustainable remediation approach might be less than if remediation options and their impacts are also actively considered as part of the land use and overall project planning stages. Earlier consideration of the sustainability impacts of remediation may provide important additional sustainability gains.

For example, the trigger for Risk Management on a contaminated site may be that the site use is to change from derelict (where risks were managed by institutional controls) to a new shopping development. This trigger sets in motion a chain of Risk Management decisions related to that land use. Let us say that for a particular development the financial model dictates a need for basement car parking, this has the consequence that the remediation interventions required will likely be ex situ. However, the contaminated material is not a groundwater threat and while it has elevated levels of contaminants, would not pose a threat unless direct contact took place. Sustainability might be an important criterion in choosing between available remediation approaches such as off site removal, on site soil washing or on site biopiling, favouring a biopiling approach in terms of waste generation, use of resources, emissions, impacts on site neighbours, costs and other factors, *for that particular site*.

However, if remediation impacts had been able to be considered at an earlier stage of design making, the "improvement" in sustainability might have been greater if for example undercroft car parking had been an option, and the contamination managed by containment and in situ bioremediation, simply because soil excavation was avoided. Fig. 20.5 illustrates this example scenario.

The question of when and how the sustainability consequences of these decisions should be considered introduces the concept of a "framework" where there may be multiple points at which sustainability is considered.

20.2.4 Frameworks

The sustainability of remediation is the consequence of a series of decisions that lead up to a remediation project, and decisions made about what remediation method is to be used within the scope imposed by these preceding decisions. It is important to be aware that this raises two sets of questions: when and how should sustainability

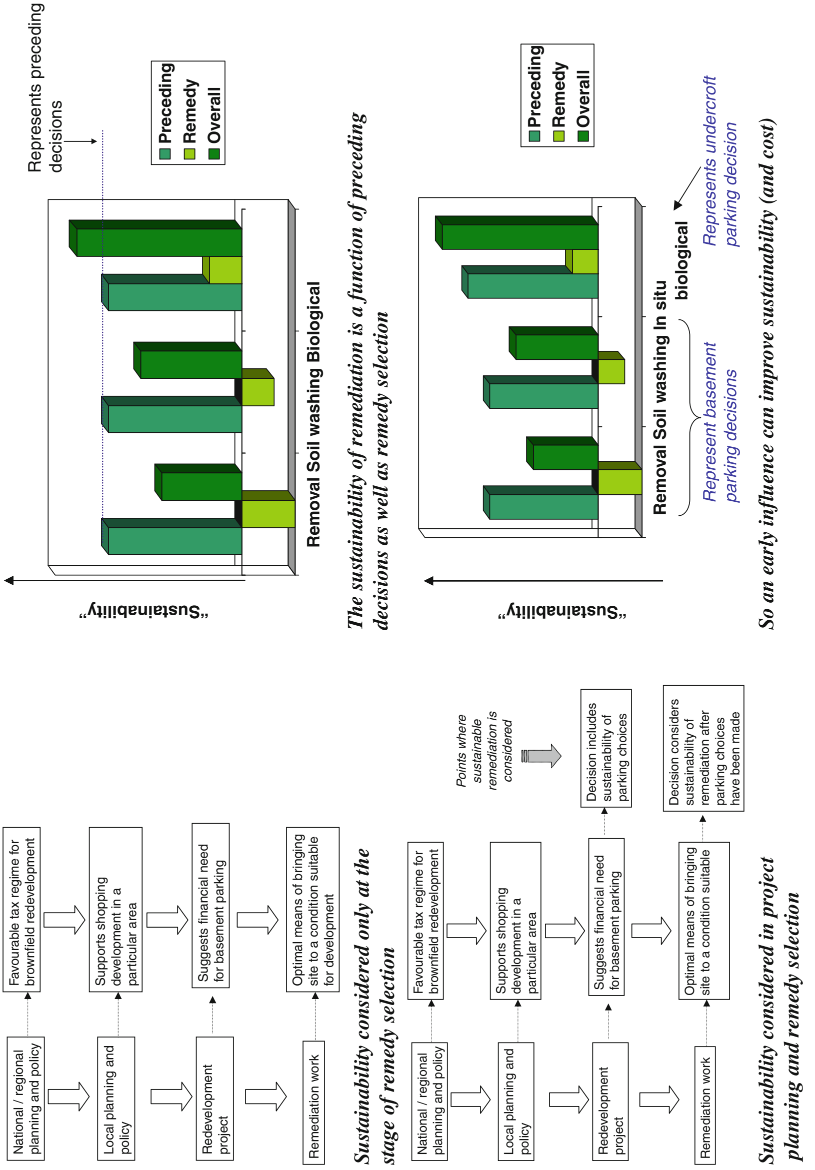


Fig. 20.5 Example scenario for sustainable remediation decision making

be considered in this decision making process?; and what is the most appropriate stage for decision making in terms of overall sustainability benefits?.

Setting out a "framework" for the decision-making process that affects contaminated land management provides some answers to these questions. The SuRF-UK (Sustainable Remediation Forum in the UK) initiative has identified several points at which sustainability based decision making is merited, which impact on contaminated land management for a particular site:[7]

- High level decision making for regional spatial planning and policy – by national/regional agencies.
- Local level land use planning and policy, by local authorities.
- Project based decision making that sets Remedial objectives (e.g. related to Risk Management/development needs).
- Remedy selection and implementation, including monitoring and verification implications.

The higher the level of decision making the greater the range of sustainability issues that are likely to be considered. For example, concerns at a regional level are likely to be with the implementation of European Directives and national legislation, such as housing and infrastructure targets, river basin management and global goals for prosperity and social equity. Regional policy will set out broad planning visions to highlight regional diversity and local distinctiveness, including amongst other things, how many houses and how much employment land are needed. At a local level decisions will focus on environmental, economic and social concerns across an area of in the order of 10–1,000 km^2. Considerations might include the zoning of areas for different types of development and use, taking into account issues such as transportation and infrastructure, flood risks and former land use, including any local Brownfield strategies. At a project level, decision making will centre on project viability, which will mean compliance with sustainability measures. These sustainability measures may be dictated by local, regional and national policies and regulation, for example sustainable construction, reducing fossil fuel dependency, re-use of Brownfield land, complying with requirements to provide amenity and links to local transportation, sustainable urban drainage and many more. At the level of remedy selection, sustainability considerations are solely directed at a choice between available remediation options.

What is apparent from this structure is that at the higher levels of decision making indicators may be fairly abstract from the point of view of remedy selection (for example relating to employment, educational achievement, prosperity, and broad environmental concerns at a regional level). In addition, the background and expertise of those taking decisions is likely to be distant from those of contaminated land management professionals. Moving down towards local level and site based decision making, indicators of sustainability will become less abstract from the point of view of remediation practioners. However, those taking the decisions are still

---

[7] www.claire.co.uk/surfuk

unlikely to have knowledge of contaminated land management in any detail, and they will be considering a wide range of sustainability issues.

Decisions taken at these "higher" levels limit the range of remediation approaches that can be undertaken, and the impacts of these limitations may be substantial. Hence a sustainable remediation framework needs to consider and influence more than simply remedy selection once all preceding decisions have been made. It needs to create an opportunity for influence at least at site management planning and local area planning, where major opportunities for improving sustainability exists. This opportunity not only includes the example above of avoiding unnecessary remediation interventions, but also linkage of remediation with other sustainable development opportunities such as renewable energy (as illustrated in the "Sanergy" case study in Section 20.5.5). A corollary of this engagement is the need for contaminated site management professionals to be able to deal with a wide range of interests and stakeholder interests, and that they can demonstrate the importance of the opportunities provided by sustainable remediation (see Section 20.2.5).

SuRF-UK has taken the view that from a pragmatic point of view for its sustainable remediation framework there are two key stages in making sustainable remediation decisions (CL:AIRE 2010):

- during the project design phase;
- at point of implementing the remediation aspects of the design.

Sustainable remediation considerations should be an influence on strategic decision making at local and regional level policy frameworks, but this is not explicitly addressed in the SuRF-UK framework which has focussed on decision making from a site or project level, illustrated in Fig. 20.6.

The SuRF-UK framework distinguishes project design decisions from decisions connected solely with remedy selection. It suggests that there is a point of no return or "glass ceiling" after which site or project based decisions cannot easily be revisited, for example because a planning or regulatory approval has been received. Beyond this point sustainable remediation decision making is necessarily limited to remedy selection that dovetails with the project design.

Even if decision making is confined to remedy selection, sustainability is an important consideration, for the following reasons:

- The cumulative effect of individual project and site based remedy selection decisions may well be large at local and regional scales, and the sustainability appraisal used for planning policy decisions at these scales is unlikely to take contaminated site management opportunities into account in a substantive way; and
- At a development project scale, decisions about the nature of construction of a built development are likely to have substantial sustainability implications.

---

[8] Figures 20.6 and 20.7 are taken from the SuRF-UK "Open Forum" Meeting, 18th March 2009, available from http://www.claire.co.uk/index.php?option=com_content&task=view&id=182&Itemid=78&limit=1&limitstart=6. These figures were redrawn in CL:AIRE 2010

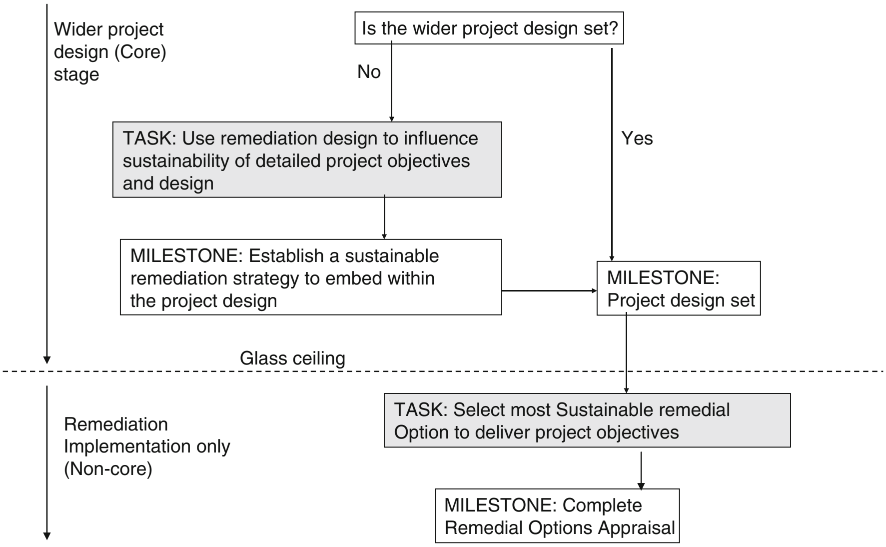


Fig. 20.6 Generic SuRF-UK framework[8]

However, that is not a good reason to ignore the "additional" sustainability benefit that might accrue from optimising the site remediation approach within the scope set for it.

A possible concern is that considering sustainability in remediation adds an unwelcome degree of complexity to decision making, and could add costs to remediation projects. Clearly, it is important not to burden a project in planning with an unreasonable information collection requirement. The logical way forward is to take a tiered or stepwise approach to sustainability analysis, beginning with simple qualitative assessments, and only moving to more detailed assessments where a simpler approach has not been able to yield clear decision support ("simple when possible, complex when necessary").

A tiered approach is essentially that the cheapest and least complicated approaches to sustainability appraisal should be used in the first instance, with more complicated appraisals only being used where these are unable to deliver a clear finding. This is consistent with a series of key requirements for sustainability appraisal tool development identified by Therivel (2004), including the need for:

- An appropriate point of trade-off comprehensiveness, rigour, transparency, user friendliness and costs. The choice should depend on the decision that the tool is informing.
- Efficiency in the tools used: the amount of time and effort they need as input should be proportional to the benefits that they provide as output.
- Multi-purpose tools, i.e. tools that can be used for several different functions and tools that bring together different disciplines; and

- Two-stage tools or processes, with a "shallow" initial stage which gives a broad-brush analysis of a problem, and a "deep" focus on those issues that were identified in the first stage as being particularly problematic, contentious or important to the decision making process.

A tiered approach has significant merit. It is flexible and provides options to assess sites to the degree necessary. An example of a tiered approach might be:

- Tier 1: Qualitative accessible approach (broad scope, rapid and low cost).
- Tier 2: Consensus / consultation to elaborate qualitative assessment.
- Tier 3: Quantitative assessment where stakeholders are unable to agree findings for a qualitative approach. In this case the preceding stages can focus quantitative work on the issues of greatest contention.

A tiered approach also supports a sustainable use of resources for the sustainability appraisal process itself.

Figure 20.7 illustrates the SuRF-UK framework's tiered approach. The SuRF-UK view is that sustainability appraisal can apply at more than one level of decision making, e.g. project design, or remedy selection. At whatever level sustainability appraisal is applied, a tiered approach should be used, beginning with the simplest (and cheapest) methods, to ensure decisions are made on the simplest basis.

The use of this tiered approach is obviously affected by the preferences of those involved in decision making. Particular stakeholders may see the need for an approach based on using scoring and weighting, or an approach linked to some form of quantification such as cost benefit analysis (see Section 20.3.4). The drivers for this may be regulatory, because of corporate preferences for quantitative indicators, or to support particular stakeholder communication efforts. For example, where a project requires a large financial investment, a formal cost benefit analysis may be required. It may make sense to integrate this with processes of sustainability appraisal, for example using a hierarchy of sustainability indicators (see Section 20.3.1) to identify the scope of the cost benefit analysis. It is possible that some

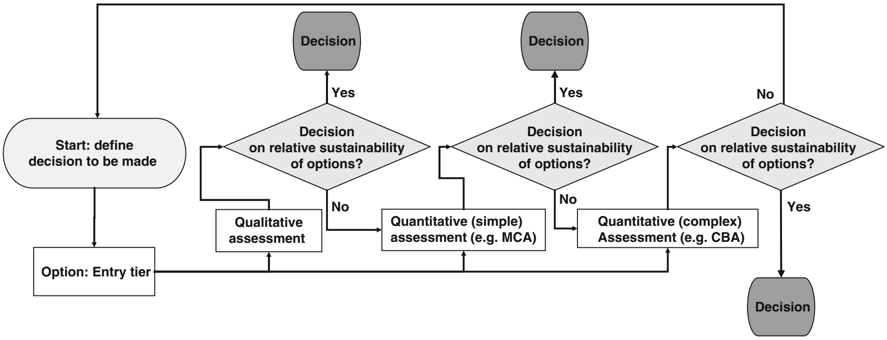


Fig. 20.7 The SuRF-UK tiered approach to sustainability appraisal

stakeholders will require some form of quantification of impacts to support a remediation choice that they do not agree with on the basis of qualitative sustainability appraisal. The development of site-specific metrics can be undertaken through a stakeholder process such as the Sustainability Assessment Tool (SAT) proposed by RESCUE.[9] Site-specific indicators or metrics will change from site to site as will the relative weighing factors for individual metrics.

Reliance solely on quantitative methods may be flawed for a number of reasons.

- All practical applications of quantitative methods are limited in scope (e.g. Table 20.1) and are not holistic sustainability appraisals. Proponents of cost benefit analysis would suggest that it has a theoretical potential to be wide ranging, however it is not clear if the valuations used will withstand scrutiny by all stakeholders (see below). Wide ranging sustainability appraisals may use techniques such as decision tables and multi-criteria analysis, but these are not strictly quantitative, but relay on scoring, ranking and weighting (see Section 20.3.4).
- There is no real means of confirming if the right metrics are being used for a particular set of circumstances beyond the rationale for a "standard" method, so the quantification may give a false impression of sustainability overall.
- There can be problems in the valuations used in quantitative techniques, such as cost benefit analysis and life cycle assessment, which are subject to uncertainties, assumptions and subjective choices which may not be transparent to all users. Valuations of human life in cost benefit appraisals can be particularly contentious.
- Some stakeholders may feel excluded by the use of quantitative techniques, for example because they do not support the valuation approach, because the metrics used do not match their views of key issues, or simply because it is seen as too "technical".

A tiered approach allows the quantitative evaluations made to be benchmarked by wider and more inclusive sustainability appraisal approaches.

20.2.5 International Initiatives in Sustainable Remediation

One of the first international initiatives to explicitly consider sustainability in remediation was the European project CLARINET: the Contaminated Land Rehabilitation Network for Environmental Technologies in Europe. CLARINET was a "Concerted Action" of the European Commission's Environment and Climate Research and Development Programme (1998–2001). CLARINET concluded that using risk based decision making in contaminated land management was entirely in line with sustainable development, and also recommended that where possible the "natural capacities" of soil and water should be used to effect Risk Management. Risk Management provides a scientific rationale for the costs of remediation that

---

[9] www.rescue-europe.com

society has to bear; the exploitation of Natural Attenuation limits those costs. However, CLARINET also found that this overarching philosophy did not mean that all remediation projects are necessarily sustainable development. CLARINET suggested that considering the true contribution of remediation work to sustainable development is an emerging challenge at least as great in its difficulty as the development of risk based decision making, and with the same capacity to profoundly change how we manage contaminated land in the future (Bardos et al. 2002; Vegter et al. 2002).

Interestingly, the same point of view was emerging in the industry community, both amongst site managers and the service providers in NICOLE (the Network for Industrially Contaminated Land in Europe).[10] NICOLE held a workshop in Barcelona, Spain in 2003 on the "Management of Contaminated Land towards a Sustainable Future: Opportunities, Challenges and Barriers for the Sustainable Management of Contaminated Land in Europe" (Bardos 2003). This meeting concluded that the meanings ascribed to terms such as "sustainable" or "sustainable development" vary widely. It also concluded that there was no common language for discussing contaminated site management in the context of sustainable development. The meeting view was that "without clear definitions everybody can claim that they are acting sustainably when sometimes perhaps they are not". NICOLE decided that it would be both a major challenge, and also a major achievement, for NICOLE to catalyse the development of a common framework, widely used across Europe in the same way that risk based decision making has become commonly used.

More recently, several initiatives have begun to address this challenge. A *Sustainable Remediation Forum* was established several years ago in North America,[11] which has recently produced a "White Paper" where its members give their review of the current state of the art (SuRF 2009). A *Sustainable Remediation Forum in the UK*[12] (SuRF-UK) has also been set up and a "Sustainable remediation working group" established by NICOLE.[13] A SURF-Australia has also been recently established,[14] and ASTM international have established a standards development committee to develop a sustainable remediation standard.[15] Their work has begun a process that will likely lead to a clearer picture of what constitutes "sustainable remediation" in Europe. The importance of this work has been highlighted by the inclusion of the following definition of remediation in the February 2009 draft text for the emerging European Soil Framework Directive: "When deciding on the appropriate remediation actions, Member States shall give due consideration to social, economic and environmental impacts, cost-effectiveness and technical

---

[10] NICOLE also began as an EU funded project, but is now a self-funding network www.nicole.org

[11] www.sustainableremediation.org

[12] www.claire.co.uk/surfuk

[13] http://www.nicole.org/WorkingGroups/WGSustainableRemediation/default.aspx

[14] http://www.crccare.com/working_with_industry/surf.html

[15] ASTM International, originally known as the American Society for Testing and Materials (ASTM), www.astm.org

feasibility of the actions envisaged". In parallel, the US EPA has established a *Green Remediation initiative* which focuses on achieving aspects of environmental sustainability in remediation.[16]

The common requirement for sustainability appraisal amongst these international initiatives is to support the adoption of more sustainable approaches for remediation projects. The predominant interest for achieving this is in providing for better project design and remedy selection. However, there are some important additional functionalities for sustainability appraisal from a business or corporate point of view that are emerging from these discussions: benchmarking technologies and providing a means for organisations to report sustainability progress across a number of projects. For example, the recent EURODEMO project[17] has suggested the use of energy efficiency as a means comparing different remediation technologies (EURODEMO 2007a, b). While the majority of this chapter focuses on sustainability from a project based point of view, Section 20.3.3 discusses these other appraisal interests.

Nathanail (Chapter 25 of this book) places remediation in the broader context of sustainable Brownfield regeneration, following on from the work of CABERNET[18] and RESCUE[19] and shows that an integrated approach to project formulation is indeed possible.

20.2.6 Communicating Sustainability and Risk Management

Risk-based site management has found its way into soil or environmental protection policy frameworks in many countries in the world. The fitness for use principle for remediation of contaminated sites supports cost-effective ways of returning contaminated sites to beneficial use. However, where there is a number of stakeholders there is usually a corresponding number of opinions on how to come to the most cost-effective and sustainable way of dealing with a contaminated site.

The engagement of stakeholders is important in sustainable decision making for three reasons. Firstly, stakeholder opinions can be an important source of information about particular aspects of sustainability (Therivel 2004). A stakeholder can be viewed (1) as any party that can affect a decision, and (2) any party that may be affected by the decision. Some stakeholders may be directly involved in decision making (for example the site owner and regulator); others may not have a direct involvement but may still be influential (for example local community interests). Secondly, inclusive decision making processes improve the robustness of decisions by widening the decision making consensus and so reducing the possibility that decisions will need to be revisited because of objection in the future. Thirdly, inclusive decision making is seen as part of good governance, which may be explicitly

---

[16] http://clu-in.org/greenremediation/

[17] http://www.eurodemo.info/

[18] www.cabernet.org.uk

[19] www.rescue-europe.com

included in the national sustainable development policy, for example in the UK (Department for Environment Food and Rural Affairs 2005a, b) and, at a European level, is included in the Public Participation Directive.

Sometimes different opinions within a group of stakeholders on how to remediate are evident from the start of the process of remediation. Sometimes, different points of view only become apparent in the course of the reclamation project. Decision analysis tools are often used to help provide common ground in these efforts. Decision making requires a shared responsibility of stakeholders (Bardos et al. 2002). Therefore communication about risks triggering remedial action is a crucial aspect in the process of sustainable restoration of contaminated sites. This communication needs to take place with all stakeholders.

There have been a number of theoretical studies on risk communication (e.g. Orr 2001; Petts et al. 2003). Experiences from consultants and problem-owners show that there are a lot of successful examples of risk communication from which important lessons can be learnt (e.g. Hazebrouck and Ledrans 2005; NICOLE 1999). The key issues for the communication on risk in contaminated land management are:

- Both technical assessments of risk and perceptions of risk need to be addressed. A good understanding of both technically identified risk and perceived risk requires good communication between experts and other stakeholders. The risks that need to be considered may well encompass risks/impacts from the remediation project itself that are of concern to stakeholders, for example: cutting trees, odour, noise or heavy traffic, which can generate objections to a remediation project.
- If the discussion becomes emotional, issues far beyond technical details of the land contamination might govern the eventual outcome, or the dialogue between stakeholders may simply break down.
- Sometimes concerns about financial disadvantage, for example for householders on affected sites, or other issues might determine the process of decision making, and other issues are given less weight.
- If stakeholder groups struggle to converge on acceptable strategies, separating the decisions about what sustainability appraisal tool, from the selection of which indicators or metrics need to be considered can simplify finding practical and acceptable compromises.
- Pro-activeness pays off. Communication about the contamination problem and possible remediation measures in advance creates trust. In dynamic situations, the availability of experts who answer questions on demand helps to keep a dialogue open and avoid critical situations. However, for this strategy to be successful it is important that the experts can communicate in clear terms and avoid the use of jargon when talking with lay people. Visualisations such as figures and diagrams can be particularly helpful, as can a shared glossary of technical terms, where these are unavoidable.

Risk-based strategies can have important benefits if the approach is broadly accepted by all stakeholders. In order to achieve this acceptance, the strategy should be the result of a process characterised by openness as well as respectful and responsible participation of all stakeholders.

The stakeholders at the centre of decision making are generally the project team, comprising the site owner and/or polluter, whoever is being affected by the contaminated site, the service provider, the regulator and planner. However, other stakeholders can be influential, such as those who might use the site (workers, possibly unions, and other visitors); those who have a financial involvement in the site or the site's ownership (e.g. banks, founders, lenders, insurers); the site's neighbours (adjacent owners and tenants, local communities and councils); and particularly for more complicated problems other technical specialists, researchers, NGOs and pressure groups.

The more complex the site (both from a technical point of view and in the context of its local circumstances) the greater will be the influence of these other stakeholders. Their input can be managed via some form of a project advisory group or board. After the decisions have been made this group can be used as a monitoring group. There is an important connection between sustainable remediation and the interests of this wider community of stakeholders, as their interests may be strongly related to the wider benefits and impacts of the remediation project being carried out. A well structured approach to sustainability appraisal can be used as a platform for positive engagement of these stakeholders and a structure for discussion and decision making that provides an objective, as opposed to an emotional, context.

It is generally beneficial to involve key stakeholders from the beginning of a project, particularly for complex or otherwise contentious remediation projects. Generally, projects tend to fall into investigation and assessment and options appraisal; remedial design; execution and verification phases, sometimes followed by a long term monitoring and maintenance phase. Execution is preceded by a remediation plan being submitted to and approved by the competent authority. Although the organisational structure during the investigation/assessment phase can differ from the execution phase, the project manager should always ensure that each stakeholder involved plays a role in the entire process of risk based site management. Proactive engagement of relevant stakeholders in all phases of the process can help minimise complaints afterwards about the violation of agreements. The problem owner should ensure, perhaps by delegation through contractors, that stakeholders are heard and that the objectives of the site remediation are safeguarded as it was agreed upon with the competent authorities and stakeholders. Periodical meetings can be used to evaluate the progress of the remediation works together with the stakeholders and to check whether there is still consensus on the path to follow.

20.3 Using Sustainability Appraisal in Remediation Option Appraisal

Remediation option appraisal forms a part of contaminated site Risk Management which can occur during project design and/or remedy selection. Assessments of sustainability are also seen as a means of differentiating remediation technologies from a marketing or promotional point of view, or for reporting sustainability performance at a corporate level across a number of projects.

20.3.1 *The Scope of Sustainability Appraisal as a Decision Support Process in Projects*

Sustainability appraisal during project design and remedy selection is typically applied in a comparative sense, i.e. which option from more than one is the "best" solution. Sustainability appraisal is applied as a *decision support tool*. Decision support exists to help those who have to take decisions cope with the complex and wide-ranging information involved in contaminated site management. Decision support can be provided as written guidance, e.g. flow sheets, or model procedures; and/or software. It aims not only to facilitate decision making, but also helps to ensure that the process is transparent, documented, reproducible and hopefully robust, providing a coherent framework to explore the options available (Bardos et al. 2002). However, decision support does not *make* a decision. That can only be done by the people involved. In Section 20.2.3 it was explained how sustainability appraisal can be used as a consideration in decision making at a number of points in contaminated site management projects, and indeed in setting spatial planning policies at local, regional and national levels that affect the overall way in which contaminated sites may be re-used.

The scope of sustainability appraisal, in terms of the number of individual factors or indicators to be considered is broad. SuRF-UK has recently assessed a wide range of indicator sets to establish how well they relate to contaminated site management, and how broadly they cover sustainable development (CL:AIRE 2009), following earlier sustainable remediation review work (Environment Agency 2000). Table 20.4 lists a series of categories within the three elements of sustainability (economic, social and environmental) that SuRF-UK has used to assess the range of coverage of different sustainability indicator sets. Each category or

Table 20.4 Headline indicator categories; table used by SuRF-UK (CL:AIRE 2010)

Environmental	Social	Economic
Impacts on air (including climate change)	Impacts on human health and safety	Direct economic costs and benefits
Impacts on soil	Ethical and equity considerations	Indirect economic costs and benefits
Impacts on water	Impacts on neighbourhoods or regions	Employment and human capital gain
Impacts on ecology	Community involvement and community satisfaction	Gearing
Use of natural resources and generation of wastes	Compliance with policy objectives and strategies	Life-span and "project risks"
Intrusiveness	Uncertainty and evidence	Project flexibility

"headline" encompasses a range of related individual indicators. The SuRF-UK review (CL:AIRE 2009) did not identify an indicator set that could already be used unaltered for sustainability appraisal for remediation.

It had been suggested during SuRF-UK and NICOLE meetings that the GRI indicator sets discussed in Section 20.2.1.2 (Table 20.2) could be considered for use in contaminated land management sustainability appraisal, as they are a comprehensive internationally agreed indicator set already in use for corporate reporting. However, even the GRI set is incomplete for remediation appraisal. For example, they do not consider soil and geotechnical functionality, landscape, the built environment and protection of archaeological artefacts. These international indicator sets may also include factors that may well be seen as irrelevant for a remediation project, e.g. "Education – Education Level – Children Reaching Grade 5 of Primary Education" *and* "Adult Secondary Education Achievement Level" from the 2001 UN indicator set. The SuRF-UK initiative is currently considering the need for and possible contents of a checklist of indicators for the consistent reporting and assessment of sustainability for remediation, based on earlier work which suggested a range of environmental indicators (Environment Agency 2000).

Two other recent UK reports have reviewed in excess of 100 individual sustainability appraisal tools from a wide range of developers and suppliers (Building Research Establishment 2004; Therivel 2004). None of these tools had a holistic coverage of the scope of sustainable development. Therefore, there is perhaps a need for further development to focus on more holistic appraisals.

An emerging concern is that the wide scope of sustainable development themes will result in an unmanageably large number of individual indicator assessments that will lead to expensive and unattractive sustainability appraisal approaches. There are two possible approaches to deal with this problem. The first is to make some kind of selection of Key Performance Indicators (KPIs). The second is to structure the consideration of indicators in a hierarchy, and use this as a means of simplifying assessments.

20.3.1.1 Using an Indicator Hierarchy

A hierarchy structures indicators in a way that reflects their inherent relationships. This type of hierarchical approach is also consistent with the tiered approach discussed in Section 20.2.3. For example, environmental indicators could be structured as set out in Fig. 20.8. A simple sustainability appraisal might simply be based on a solicitation of stakeholders views of the key environmental, economic and social impacts of particular remediation options as routinely happens under the English planning system. This might be perfectly adequate for small remediation projects in less contentious areas, where the decision level is solely connected with remedy selection. The next tier might be to base a sustainability appraisal on a qualitative assessment of "headline" indicators which incorporate a range of individual indicators. For example, Table 20.4 provides 18 such headlines. This type of qualitative appraisal may be sufficient. If it does not provide a clear decision making rationale,

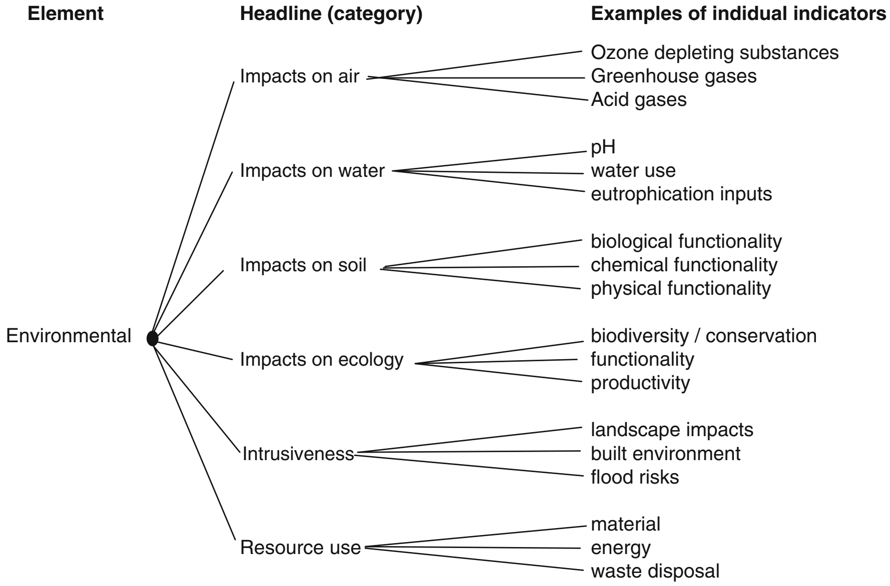


Fig. 20.8 Example indicator hierarchy for the environmental element of sustainable development

then individual indicator assessments may need to be made, but presumably only within those headline categories for which consensus could not be reached.

Double counting is a potentially serious issue that gives undue weight to an issue if it is represented by more than one criterion. However, in attempting to avoid double counting, there is a danger of oversimplification, and that key issues that resonate with stakeholders are lost (Mitchell 2005), as discussed in Section 20.2.1.3. Grouping indicators under headlines balances the need to avoid double counting with the need to avoid over-simplification.

The importance of this kind of structured approach to considering indicators reflects the functionality of project based decision making, where reaching consensus between different stakeholders for a particular site. Sustainability appraisal is specific to the circumstances of a particular site or project. It depends on the specific environmental, economic and social context of the site being considered, and also the people and organisations involved. The reason for carrying out sustainability appraisal will be related to gaining some form of agreement from planning or regulatory authorities or a local community, or some form of reassurance from a corporate point of view that the site owner can defend their decision as sustainable, or some combination of similar reasons. Many factors important in sustainability are likely to be viewed subjectively by different stakeholders, for example impact of a project on a historic built environment. Hence, the only *real* decision will be that the different stakeholders involved *agree that for a particular criterion the most sustainable approach is "x"*. Therefore, from a pragmatic point of view, sustainability appraisal only needs to go as far as is necessary to reach this consensus, and the sustainability appraisal process should be a tool for arriving at a consensus based decision. This is not just because inclusive decision making may actually seen as fundamental to sustainable development, but also because it makes for a more robust decision that is less likely to be challenged.

20.3.1.2 Using Key Performance Indicators (KPIs)

Advantages that have been suggested for making a *limited* selection of KPIs for assessing sustainability in remediation is that they could be connected to wider corporate sustainability reporting, or the use of indicators that are being used to promote the advantages of particular remediation approaches (see Section 20.3.2). The KPIs suggested are almost always limited to environmental parameters, estimates of Risk Management performance and measurements of direct cost.

The difficulty with selecting particular indicators as Key Performance Indicators (KPIs) for project planning and remedy selection is that the selection will necessarily be partial. The selection is likely to be based on perceived importance of particular issues: e.g. carbon balance might be perceived as more important than soil functionality. This implicit weighting allows plenty of space for disagreement and controversy about the sustainability assessment approach let alone its findings. The smaller the set of KPIs is, the greater is the possibility that an influential component of the true overall sustainability will be ignored.

KPIs can also be limited to quantifiable measures which, it is felt, remove the subjectivity inherent in qualitative approaches. However, quantitative assessments have their own limitations, summarised in Section 20.2.3, which in turn limit their usefulness in sustainability appraisal.

20.3.1.3 Agreeing Sustainability Indicator Approaches for Remediation

The US EPA Green Remediation initiative uses a series of measurable environmental “indicators”. SURF in the USA takes a view of sustainability across all three elements (environmental, economic and social). This appears also to be the approach in Europe. Indeed the draft Soil Framework Directive was explicit about environmental, economic and social indicators. At present, there is no European consensus on which indicators might be used for the assignment of the sustainability of remediation. There is considerable debate about the use of a limited set of quantifiable KPIs versus other approaches, particularly within the NICOLE Sustainable Remediation Working Group.

The view of the chapter authors is that in many cases qualitative approaches will either be sufficient for project based decision making, or to identify what the stakeholders involved consider to be the key sustainability “drivers” for a project (see Section 20.4) that need to be quantified. These drivers are likely to include a component that is strongly project specific. Where some or all of the stakeholders involved in a project decision prefer a *general* quantitative approach, it may still be prudent to carry out a wide ranging qualitative sustainability appraisal as a first step, to ensure consideration of the wider sustainability outlook.

20.3.2 Using Sustainability for Technology Promotion and for Corporate Reporting

20.3.2.1 Promotion of Remediation Technologies

An attractive proposition for some technology vendors and service providers might be to have some kind of a benchmark which shows that their technology is more sustainable than someone elses, analogous to how household appliances are now commonly sold with energy ratings. There is no fundamental measure of sustainability that can be applied to show that a particular remediation technology will be better from a sustainable development perspective whatever the remediation project context. However, there may be aspects of a particular technology which may be regarded as beneficial, particularly from the point of view of controlling environmental impacts. These may be labelled "Key Performance Indicators" by those keen to promote them. There is also much interest in the use of life cycle based tools to provide generic measures of environmental technology performance, particularly in the waste and recycling management sector (Bender et al. 1998; Dalemo and Oostra 1997, Environment Agency 2009; Wittmaier et al. 2009). Current policies related to climate change also create an interest in the carbon intensity or energy efficiency of particular remediation approaches (EURODEMO 2007b). Caution is needed in applying Life Cycle Assessments in a generic way, rather than in a comparative way (Environment Agency 2000; Finnveden 2000). The benefits being promoted will undoubtedly be selective, because many aspects of sustainability may not be readily quantifiable. Furthermore, sustainability is a function of a *whole* project and how a technology is used, as illustrated in the case studies below. For example, a particular remediation technology may contribute to a project that is sustainable in an overall sense, even if, for example, its carbon intensity is greater than that of a competing technology. Hence, while it is perfectly legitimate to promote particular environmental advantages of a remediation technology, the decision on using that technology for a particular project has to be based on a site specific sustainability appraisal if "sustainable remediation" is to be achieved. Nonetheless, technology vendors able to provide context-relevant data on sustainability impacts (across all three elements of sustainability) to facilitate sustainability appraisal may be better able to compete in a sustainable remediation market.

20.3.2.2 Linkage to Corporate Reporting

The broad ways in which corporate reporting on sustainable remediation may be carried out include finding some means of overarching sustainability appraisal across all projects, and demonstrating that individual projects have met their specific sustainability criteria based on the consensus across the stakeholders involved in the decision making process. The dominant sustainable remediation function is ensuring that sustainable projects are carried out, and this decision making will be project specific. This makes the use of an all encompassing set of corporate sustainability metrics across all projects difficult to achieve. Furthermore, corporate sustainability reporting for an organisation may be across a wide range of operations, not just contaminated site remediation. Consequently, issues of interest in remediation such as soil functionality or landscape conservation may not form part of general corporate reporting, particularly if these are linked to a limited set of quantified organisational goals such as carbon, water and waste management.

One way forward from this dilemma is to focus on a set of corporate "Key Performance Indicators" which reflect the organisations overall sustainability reporting. These will only be a partial view of remediation sustainability, but these indicators could be supported by reporting of how individual projects have met specific sustainability criteria agreed by the stakeholders involved in each project.

20.3.3 Frameworks and Boundaries

For any comparison to be valid like must be compared with like. For example, comparing an on site biopile with removal to landfill will not be valid unless the boundaries of the system being compared are the same. It is important to define objectives as closely as possible, being specific, for example, about the exact scope, and the exact options being considered. The better defined the scope and options, the more reliable the sustainability appraisal. There are four broad categories of boundaries that should be considered:

- the "system boundary", which is the boundary affected by the framework within which contaminated site decision making is made and includes the scope for remediation set by preceding management decisions;
- the "life cycle assessment boundary" – while Life Cycle Assessment focuses only on a range of environmental impacts, "Life Cycle Thinking" may also be appropriate for a wider range of sustainability indicators (Koneczny et al. 2007);
- geographical boundaries; and
- the duration over which effects are to be considered.

The system boundary describes the "edges" of the system being considered, i.e. where it interfaces with the surrounding environment, society or economic processes or other systems. The system boundary (US Environmental Protection Agency 2006) encompasses the following:

- Scope of the system being considered: for example a choice of *ex situ* remediation/disposal methods for dealing with excavated soil generated during construction of a basement, some of which may be contaminated. The scope of what is being considered is determined directly by the overall framework within which contaminated site management decisions have been taken as described in Section 20.2.3.
- Operations such as: conceptualisation, design, mobilisation, delivery, construction, utilisation, production, refurbishment, maintenance and decommissioning.

A rule of thumb is that the system boundaries should be the smallest set of project operations that allows options to be compared across the same system for the options being considered (i.e. like is compared with like) and fully describes likely benefits and impacts.

Another important system consideration is that any comparison is made over the same scale of operation. For example, it would not be helpful to make a direct unscaled comparison between a pilot scale bioremediation process and full scale excavation and disposal operation. Many sustainability indicators would be completely different on the basis of the relative scale of the operations alone, for instance considering the impacts of vehicle movements. Quantitative Life Cycle Assessment and carbon foot-printing uses a concept called a "functional unit" (Carbon Trust et al. 2008a) to ensure that like is compared with like, for example: "An appropriate functional unit for an ex situ treatment processes might be the treatment of a specified amount of contaminated soil (e.g. per tonne of treatment capacity per year)".

Life cycle assessment boundaries consider how far the option being considered should be broken down into sub-units requiring some sort of analysis. A key part of understanding life cycle boundaries is the concept of *cradle to grave.* The cradle is the origin of materials or substances being considered. The grave is their ultimate fate (US Environmental Protection Agency 2006). Life cycle comparisons consider effect from origin to fate, i.e. cradle to grave. There are practical limits to how far a cradle to grave comparison is feasible for a sustainability appraisal process. For example, at its most extreme a determination might consider amongst other things the impacts of using/making resources in each individual pipe, joint and fan used in a soil venting process, even each nut and bolt. Clearly this level of detail poses great demands on an assessor's time and resources. Another problem is how to consider equipment that might be used over several projects, for example an excavator or other plant. The total number of projects each of these vehicles will be used for is probably unknown. Hence, it becomes very difficult to apportion impacts to one single project for the *production* of components that are used multiple times. To set life cycle boundaries stakeholders will need to agree:

- what is a practical limit to the level of detail, for example whether staff welfare consumables and facilities such as protective equipment and toilets should be considered or disregarded as broadly similar for each option being compared;
- whether "generic" assessments for particular universal components such as hardware can be accepted;
- what is *consumed* by the project and what is used non-exclusively by the project; where an item is consumed the impacts of its production and use must be considered; where it is re-used and not consumed the impacts of its use must be considered but in many cases the impacts of its production might be discounted;
- where facilities already exist on a site, such as monitoring wells, the environmental impacts of their construction can be discounted from the sustainability appraisal; where facilities are constructed they should be regarded as *consumed* by the project.

An interesting point in this regard is where materials are reclaimed and re-used during the project. Recycling, re-use and recovery have specific policy based meanings, which will affect what can be considered for the policy-directed indicators. A further difficult point is the use of materials to fill void space. Filling void space in a landfill site is essentially consuming a waste management resource, whereas filling void space at an on-site or off-site development project may displace virgin materials or other recycled materials but does not consume a waste management resource. However, the placement of the material may carry impacts, for example risks from leachate, which would need to be considered in a sustainability appraisal. This activity may also have wider sustainability linkages. For example, it may create an opportunity to install ground source heat pumps and thereby reduce the environmental impact of the subsequent development by providing a source of renewable energy. Alternatively, it may be integrated with a "sustainable urban drainage" system, and the linkage with remediation work may also reduce the cost on implementing these additional opportunities.

The life cycle and system boundaries are different considerations. The life cycle boundaries describe how far a particular trail of consequences should be followed and to what level of detail. The system boundary describes the systems being compared: what sets of operations are being compared.

The intuitive understanding most people have of geographical boundaries is a site perimeter. However, the sustainability appraisal has to consider impacts and benefits across the system and life cycle, which may occur:

- at the sites of production or application processes;
- at supplier sites (including how a project approach might affect waste collection);
- through transportation and distribution;
- through distant impacts for example effects on air and water, or distant effects of increased traffic.

It may be useful for some appraisal purposes to distinguish between local and distant effects, or other categories such as regional, national etc. Geographical boundaries would need to be agreed to define such terms. The overall sustainability appraisal is *independent* of these classifications. However, there may be reasons to consider local and distant effects differently in selecting a remediation approach, particular with regard to social impacts.

The duration over which effects are to be considered is also an important consideration. Sustainability appraisal considers *changes* in indicators due to the remediation work being assessed. The initial time boundary, *t=o,* is the commencement of the *operations* defined by the system boundary. The remaining time boundary is of course the point in time beyond which effects are no longer considered. In absolute terms, assessments of effects based on emissions, such as of greenhouse gases or toxic substances, or effects of the disposal of wastes should be considered over very long periods. The overall sustainability appraisal is *independent* of these classifications. However, there may be reasons to consider short and long term effects differently in selecting a remediation approach, particular with regard to social impacts, with particular stakeholders may be interested in distinguishing short and long term effects. The duration boundaries are essentially the time classifications used to draw these distinctions. An important aspect of the time boundary for assessments using moneterisation is the discount rate that will be applied in net-present value calculations.

20.3.4 Techniques and Tools and Their Applicability

Table 20.1 listed a range of techniques that have been applied to sustainability appraisal, or aspects of sustainability appraisal, or have potential to be used in sustainability appraisal. This section reviews them each in more detail. None of the techniques used in sustainability appraisal is an absolute valuation. All are fundamentally flawed in that they can only assess what we know or perceive may be an issue. For example, would a contemporary sustainability appraisal of DDT in the 1930s, if such a thing existed, have taken into account its impacts on birds of prey, as biomagnification was not a known phenomenon at that time.

20.3.4.1 Systems Using Scores, Rankings, Weightings, Including Multi-Criteria Analysis

Many techniques that are qualitative, such as cost effectiveness analysis, use scores (some kind of figurative effort of the scale of an effect) and often also use weightings (some kind of figurative effort of the relative importance of an effect). Weightings and scores are typically combined, scaled and otherwise arithmetically processed, for example using Multi-Criteria Analysis (see Box 20.1). This may yield a result that appears rather definitive, perhaps more definitive than should be inferred from the original process of scoring and weighting since there are several limitations in using scores and weightings.

Types of scoring might include symbols like ++, +, 0, –, —; or numeric values such as scores from 1 to 10 or 1 to 100%. A fundamental limitation is that this scoring will always be subjective, unless it includes an element of formal quantification. Where scores are determined on the basis of best guess there is no way to guarantee that one person’s “++” is not the same as another person’s “+”, setting aside any differences in stakeholder perspective, but purely because some people tend to score high and others low. What is particularly problematic about scores is that they add a “numeric” gloss, i.e. the illusion that something is known in a quantitative way when it is not. These apparent numbers can be manipulated and processed, perhaps to the extent where their original unreliability may be forgotten. Scoring can take a lot of time for not much added value, for example, as they wrestle with questions like: “should a score be 65% or 70%?” Weightings are always by their intrinsic nature subjective, as different stakeholders can weight differently according to different perceptions of importance. Deciding weightings can add complication and contentiousness to the appraisal process, and they are in effect a valuation without formal derivation.

Box 20.1 Multi-Criteria Analysis (MCA)

A range of qualitative sustainability appraisal techniques have been published based on scoring systems, with different levels of complexity, for example for regional spatial strategies (Office of the Deputy Prime Minister 2005a), project appraisal (Ministry of Defence 2006).

Multi-Criteria Analysis (MCA) is often used in decision making. MCA is a structured system for ranking alternatives and making selections and decisions. Considerations used in MCA are: the magnitude of an effect (score) and the importance (weight). MCA describes a system of assigning scores to individual effects (e.g. impact on traffic, human health risk reduction, use of energy, etc). These can then be combined into overall aggregates on the basis of the perceived importance (weighting) of each score. With MCA, ranking and decision making processes can be made very transparent (Environment Agency 2000; Wrisberg et al. 2002).

MCA is not a tool that directly analyses physical or monetary information. Rather it is an analytical tool at a higher level, bringing together different considerations in a structured way. However, techniques such as CBA and Life Cycle Assessment apply MCA principles in their use of weightings, scoring (valuations) and aggregation, as does the sustainability appraisal described in this guidance. MCA describes a range of techniques, and at its most complex, MCA might include analyses of individual preferences of stakeholders for weightings and quantitative valuations (such as LCA techniques) for deriving scores.[20]

The use of weightings is a difficult decision for an assessor to make, as there may be strong demands for "importance" to be considered in the sustainability appraisal. However, the impact of weighting is to skew the findings of an appraisal. Consider a very big "impact", which would lead to a very high score, combined with the nature of the effect being considered to have a very high importance, which leads to a high weighting. The combination of the score and weighting is geometric (multiplied). However, the aggregation of individual score and weighting products is arithmetic (added). Consequently the sustainability appraisal will be dominated by the highest scores with greatest weightings. While this may have been the intention, the consequence of these geometric products with addition may be to submerge the effect of other factors on the sustainability appraisal to seeming irrelevance.

An alternative to using scores is to use rankings. Rankings can be evidence-based without the limitation of attaching "values". Rankings simply show which options are better than others on the basis of available evidence. However, once generated

---

[20] In this scenario MCA approaches are used both in making valuations, and combining different valuations, for example environmental impacts and costs.

they can be manipulated arithmetically. However, rankings also have limitations; in particular, they do not consider the "scale" of differences that might separate particular options: i.e. they only say Option A is better than Option B, but not that Option A is "much" better or only "slightly" better than Option B.

Rankings can also be used instead of weightings, for example, labelling indicators as "high", "medium" or "low" importance. A simple sensitivity analysis can then be used to see the effect of removing indicators of "low" importance from a sustainability appraisal.

Scaled rankings can be used to determine weights. Rather than saying "Option A is better than Option B", scaled ranking say "Option A is 3x better than Option B." By limiting the scale to 3x, 5x, 7x and 9x stakeholders are given some but not too much scope to express the strength of their preference (Nathanail 2009).

20.3.4.2 Best Available Technique (BAT)

The concept of Best Available Technique (BAT) was described on the IPPC Directive (Directive 96/61/EC).[21] The definition of BAT is "the most effective and advanced stage in the development of activities and their methods of operation which indicates the practical suitability of particular techniques for providing in principle the basis for emission limit values designed to prevent and, where that is not practicable, generally to reduce emissions and the impact on the environment as a whole". Where there is a choice, the technique that is best overall will be BAT, unless it is not an "available technique". There are two key aspects to the availability test:

- what is the balance of costs and advantages?; This means that a technique may be rejected as BAT if its costs would far outweigh its environmental benefits; and
- can the operator obtain the technique (Department for Environment Food and Rural Affairs 2007b).

BAT determinations are therefore a partial sustainability appraisal and a limited form of cost benefit analysis. BAT considers some aspects of sustainability, but historically these were not formally linked to indicators derived from the three elements of sustainable development.

20.3.4.3 Carbon Footprint ("Area")

A *carbon footprint*[22] is a measure of the impact human activities have on the environment in terms of the amount of greenhouse gases produced, measured in units of carbon dioxide. A carbon footprint is made up of the sum of two parts, the direct/primary footprint and the indirect/secondary footprint. The primary footprint is a measure of direct emissions of CO_2 from the burning of fossil fuels for example

---

[21] http://ec.europa.eu/environment/ippc/

[22] http://www.carbonfootprint.com/carbon_footprint.html

to support the energy use of a remediation process. The secondary footprint is a measure of the indirect CO_2 emissions from the whole lifecycle of products used within the remediation – those associated with their manufacture and their eventual breakdown. *Note the carbon footprint is not measured in terms of area.* The world's first standard approach was recently published in the UK (Carbon Trust et al. 2008a, b).

Carbon footprint assessments have been used to describe remediation projects, for example the proprietary tool: Remediation Options Carbon Calculator.[23] A carbon footprint appraisal is a "quantitative" measure of a system's possible impact on global warming. It may have some importance as part of a corporate carbon reporting requirement, but it is not a substitute for sustainability appraisal. It can inform about one indicator within one element (the environmental element) of a sustainability appraisal.

20.3.4.4 Carbon Balance (Flows)

Carbon balance diagrams illustrate calculations of tonnes of carbon in various inputs and outputs to an environmental management process, and how this balance changes for different scenarios (Department for Environment Food and Rural Affairs 2006b). For waste management (and so probably also remediation) the major flows of carbon/greenhouse gases and energy result from: the use of fuel and energy in processing; the transportation of materials to and from sites; direct releases from materials on processing (e.g. biological processing or thermal treatment) or disposal in landfill; avoidance of greenhouse gas emissions or energy use elsewhere in the economy; and sequestration of carbon in landfill and soil. Carbon balance diagrams show the fate of carbon for each material and scenario in detail, considering: the carbon that remains within the material fraction following treatment and re-use or disposal; carbon that is sequestered in landfill or some other soil carbon sink; carbon that is contained in products re-used off site; and carbon that is released to atmosphere, as carbon dioxide (fossil/biogenically derived) or methane. The diagrams also include greenhouse gas balance calculations shown in tonnes of equivalent carbon dioxide. A carbon balance is a detailed quantitative assessment which not only compares options, but also identifies opportunities where improvements may be made. However, as for a carbon footprint appraisal, it relates to only one factor within an overall sustainability appraisal.

20.3.4.5 Cost Benefit Analysis

Cost benefit analysis is a form of economic analysis in which costs and benefits are converted into monetary values for comparison (known as "private costs"). Cost benefit analysis considers a diverse range of impacts (known as "public costs") that may differ from one proposed solution to another, such as the effect on human

---

[23] http://atkinsrocc.com/Public/Default.aspx

health, the environment, the land use, and issues of stakeholder concern and acceptability by assigning values to each impact in common units. Deciding which impacts to include or exclude from the assessment is likely to vary on a site-by-site basis. In many instances, it is difficult to assign a strictly monetary or quantitative value to many of the impacts. Hence, assessments can involve a combination of qualitative and quantitative methods (Environment Agency 1999, 2000b). A range of valuation techniques may be used try and convert particular impacts into monetary values, as summarised in Table 20.5.

Table 20.5 Valuation techniques used in CBA and their limitations[a]

Technique	Limitations
Contingent Valuation Method (CV) is based on assessments of consumers willingness to pay (WTP) for something; or their willingness to accept (WTA) compensation for it. It is assessed by surveys of "consumers" (typically using a hypothetical scenario) who are asked to assign monetary values to both WTP and WTA	There is a considerable debate over the validity of CV in assessing environmental impacts and benefits. CV can confuse wishes with preferences. Potential errors and bias can occur in several ways using these methods, for example, resulting from the way that questions are asked, and *perceived*, by those being questioned and intrinsic differences between the hypothetical scenario and the real situation.
Hedonic Pricing (HP) is based on relationships between the levels of environmental services (e.g. noise levels) and the price of marketed goods (e.g. houses). It cannot be used to estimate the subjective factors that cannot be seen as directly affecting marketed goods.	HP is sensitive to the following errors and bias: the possibility of omission of key variables from the examined relationship, unknown correlation between "independent" variables; assuming a uniform market when it is not in reality and perhaps most importantly, HP relies on several restrictive assumptions, for example, that purchasers are fully aware of soil quality data and its implications and that the land development market is at or near to equilibrium.
Production Function Methods are similar and infer value from marketed goods and services. There are two broad approaches: avoided cost (AC): evaluation of environmental quality through quantification of averting expenditure (i.e. how much are people willing to pay to avoid or protect them from a decrease in environmental quality?) and the dose-response (DR) method, where physical effects of contamination on the environment are evaluated and used within an economic model.	AC can over-estimate the level of expenditure related to the specific environmental change of interest. A contentious area where this technique has also been applied is in valuing human life, for example in assessing the benefits of a town by-pass scheme. A key difficulty with using DR is the selection of the economic assessment model.

Hanley and Spash (1994); Mulberg (1996)

[a] Based on work originally carried out for the University of Nottingham Masters in Contaminated Land Management

Cost benefit analysis is widely used in policy and project appraisal in the private and public sectors, for example the UK *Green Book* (HM Treasury. Cost benefit analysis has been widely applied to contaminated site management decision making (Bardos 2008a, b). It is seen as having the potential to consider whatever range of indicators is seen as important for sustainability appraisal. However, it does have some serious weaknesses (Therivel 2004), which include the following: there is no standard “checklist” of indicators, so CBA is highly specific to the circumstances and method used for each particular assessment; the valuation procedures for public costs are both highly technical and also subject to serious inherent weaknesses as set out in Table 20.3. Consequently they may not be inclusive of/acceptable to all stakeholders. This problem is exacerbated where contentious “standard values” e.g. for a human life are used in a Cost benefit analysis , or values for transportation or other activities are imported into a remediation Cost benefit analysis from another analysis that may be totally unrelated even to the environmental sector let alone contaminated site management. The link between evidence and assumption may in these circumstances be rather tenuous. Some procedures include a sensitivity analysis step which allows decision-makers to question their judgements and assumptions through the eyes of other stakeholders.

20.3.4.6 Cost Effectiveness Analysis

Cost-effectiveness analysis is a simplified derivative of cost-benefit analysis used in the UK. The aim of cost-effectiveness analysis is to determine “. . . the least cost option of attaining a predefined target. . .” without a monetary measurement of benefits (Environment Agency 1999). Costs are calculated conventionally and benefits are scored individually. An aggregate score for benefits is then divided by cost to provide a measure of “cost effectiveness”. The derivation of scores is an application of multi-criteria analysis (see Box 20.1). An example applied to site remediation is given in Harbottle et al. (2008a, b).

20.3.4.7 Eco-Efficiency

Eco-efficiency (EURODEMO 2007b) is reached by the delivery of competitively priced goods and services that satisfy human needs and bring quality of life, while progressively reducing ecological impacts and resource intensity throughout the life cycle to a level at least in line with the earth’s estimated carrying capacity. Seven critical factors for eco-efficiency have been identified:

- reduce material intensity of goods and services;
- reduce energy intensity of goods and services;
- reduce toxic dispersion;
- enhance recyclability;
- maximise sustainable use of renewable resources;
- extend product durability;
- increase service intensity of goods and services.

20.3.4.8 Ecological Footprint

The conceptual idea of an ecological footprint is that it is the area of productive land and water ecosystems required to produce the resources consumed and assimilate the wastes produced (Chambers et al. 2007). The Department for Environment Food and Rural Affairs (2007c) states that it provides a measure of the extent to which human activities exceed two specific environmental limits: the availability of bioproductive land and the availability of forest areas to sequester carbon dioxide emissions. It is derived from data from resource flow analysis, about a range of activities such as transport, energy use, materials and product consumption, waste production and water use. The impacts of these activities are converted into a common currency, *global hectares* (gha). Using this common unit, a broad range of impacts can be aggregated to derive ecological footprints for products, processes, organisations, etc. The method does have limitations. It does not incorporate all aspects of ecological resources and services, and it excludes the use of non-renewable resources. Related concepts are water footprints (Waterwise 2007[24]) and waste footprints (Waste and Resources Action Programme 2007).

Like eco-efficiency, ecological footprint analysis focuses on environmental sustainability issues, and perhaps not all aspects of environmental sustainability; for example it does not consider the aesthetic value of landscapes, conservation and Biodiversity issues or the built environment. Hence it is not really adequate as a surrogate for sustainability appraisal, although it can inform certain aspects of sustainability appraisal.

20.3.4.9 Energy Intensity/Efficiency

Energy intensity is a simple metric that can be used to compare processes such as waste management processes, for example, kWh to treat a tonne of waste. Energy intensity may also be an indicator of wider environmental effects, such as greenhouse gas emissions and emissions of acidic gases. Energy efficiency has been proposed as a means of comparing the overall environmental impact of remediation technologies by the European EURODEMO project (EURODEMO 2007a, b). Similar metrics are water intensity and carbon intensity (Nichols and Looney 2007). These metrics can inform individual aspects of the environmental element of sustainability appraisal, but do not offer a complete substitute. The use of renewable energy complicates both energy intensity and carbon footprint assessments, and individual tools need to be carefully scrutinised to determine whether benefits from renewables are properly accounted for.

20.3.4.10 Risk Assessment

Risk Assessment is a way of evaluating the likelihood and magnitude of harm or detriment caused by exposure of receptors to potential hazards, for example from

---

[24] See also www.waterfootprint.org

contaminated sites. A hazard is a substance or situation, such as a contaminated site, which has the potential to cause harm, e.g., adverse human health effects (Chapters 5–12 of this book), the soil ecosystem (Chapters 13–16 of this book), groundwater (Chapters 17–19 of this book), damage to underground structures, etc.). Risk Assessment will likely have been carried out already as a part of contaminated site management and remediation option selection. Risk Assessment findings can inform aspects of sustainability appraisal linked for example to water and soil quality.

20.3.4.11 Environmental Impact Assessment/Strategic Environmental Assessment

Environmental impact assessment describes a procedure to make a structured appraisal of a broad range of environmental effects of a particular *project*. In the EU, Environmental impact assessment is subject to Directive 85/337/EEC.[25] Environmental impact assessment methods are not prescribed in detail but tend to use stages such as (Department for the Environment Transport and the Regions 1999):

- screening: narrows the application of EIA to projects that may have significant environmental impacts;
- scoping: identifies the potential environmental impacts to ensure the assessment focuses on the key issues for decision-making;
- identification: of key environmental impacts;
- consideration of alternatives: in terms of sites, designs, processes;
- prediction of impacts: predicts the magnitude of key impacts;
- evaluation of significance: assessment of significance of the key impacts;
- mitigation: proposal of measures to prevent, reduce or rectify the impacts;
- documentation: presentation of EIA results for clear communication;
- review: systematic appraisal of the quality of the environmental statement;
- post-decision monitoring: to assess the ex post effect of the project on the environment;
- post-project audit: comparison of actual outcomes with predicted outcomes to assess the quality of predictions and effectiveness of mitigation.

Strategic Environmental Assessment is a systematic decision support process aiming to ensure that environmental and possibly other sustainability aspects are considered effectively in *policy, plan and programme making* (Fischer 2007; Office of the Deputy Prime Minister 2005b). In Europe, Strategic Environmental Assessment is undertaken to meet the requirements of European Directive 2001/42/EC. Key principles in Strategic Environmental Assessment include the

---

[25] http://ec.europa.eu/environment/eia/

promotion sustainable development. A range of techniques may be used in SEA including sustainability appraisals and Cost benefit analysis.

Environmental impact assessment is only required (by legislation) for projects beyond a certain size, and not all remediation projects will trigger the need for an Environmental impact assessment in their own right. EIA could also be triggered for remediation as part of a larger development project. Environmental impact assessment does not consider the full range of factors that would be considered in a full sustainability appraisal. However, it may be beneficial to carry out sustainability appraisal if an Environmental impact assessment requirement has been triggered to provide a balanced comparison of available options.

Strategic Environmental Assessment affects national, regional and local policies that might affect contaminated site management, particularly related to spatial planning. In the UK authorities are encouraged to carry out Strategic Environmental Assessment and sustainability appraisal in parallel for spatial planning local and regional policy development (Office of the Deputy Prime Minister 2005a, b).

20.3.4.12 Financial Risk Assessment

Major projects affect businesses and administrations through their potential to influence liquidity, solvency and overall financial performance. These are financial risks relating to an organisation's ability to meet its corporate and project objectives. The precise nature and extent of financial risk depends on the context in which the project is undertaken. Financial risk relates to the internal rate of return (IRR) or Net Present Value (NPV). IRR represents the return that can be earned on the capital invested in a project; the risk is of it being reduced to a point at which a project becomes commercially non-viable. NPV represents the present day cost of some action taken at some time in the future; in essence the present day value of that distant cost is discounted by the applicable interest rate over that period of time. Financial Risk Assessment guidance for planning remediation work has been produced in the UK (Finnamore et al. 2000), and where financial Risk Assessment has been carried out it may inform aspects of the economic element of a sustainability appraisal.

20.3.4.13 Industrial Ecology

In the industrial ecology concept industrial processes are analysed as if they were living processes (industrial metabolism). Managers of the industrial system consider it, at every level, as a set of organisms, subject to ecological constraints, like any other member of an ecosystem. "Industrial metabolism traces material and energy flows from initial extraction of resources through industrial and consumer systems to the final disposal of wastes" (Lowe et al. 1997). Industrial metabolism can be used as a basis to derive "metrics" or indices of an industrial system's efficiency and productivity, for example: ratio of virgin to recycled materials, ratio of actual/potential recycled materials, ratio of renewable/fossil fuel sources, materials productivity, energy productivity, resource input per unit of end-user service. These indices may inform aspects of a sustainability appraisal, but do not substitute for sustainability appraisal.

20.3.4.14 Life Cycle Assessment

Life Cycle Assessment is a tool to evaluate the environmental consequences of products or services from cradle-to-grave, and their use (Danish Topic Centre on Waste and Resources 2006; Wrisberg et al. 2002). In the context of contaminated sites, such a function might be the remediation of a contaminated site (Shakweer and Nathanail 2003). The main features of LCA are as follows:

- LCA follows a *cradle-to-grave approach*: all processes connected with the function, from the extraction of resources until the final disposal of waste, are considered.
- LCA is *comprehensive* with respect to the environmental interventions and environmental issues considered. *In principle*,[26] all environmental issues connected with the function are specified as resulting from extractions, emissions and other physical interventions like changes in land use.
- LCA may provide *quantitative* or *qualitative* results. With quantitative results it is easier to identify problematical parts of the life-cycle and to specify what can be gained by alternative ways to fulfil the function.

LCA reports may also be accompanied by assessments of the economic value of any impacts reported (e.g. "human toxicity" and may include impacts that could be considered social rather than environmental, such as injuries at work (Koneczny and Pennington 2007). LCA based approaches have been used in several decision support tools for remediation assessment (Bardos et al. 2002). Perhaps the most widely used of these tools is the Dutch *REC* system. REC was developed in the early 1990s (NOBIS 1995a, b). The REC approach derives quantitative criteria for "risk reduction" (R), "environmental merit" (E) and "cost" (C). The "environmental merit" assessment is based on a Life Cycle Assessment related approach. The REC tool provides three indices for assessing projects: R – an index of Risk Management performance; E an index of environmental merit based on life cycle techniques considering: and C an index of direct costs. The environmental merit index considers:

- improvement of soil quality;
- improvement of groundwater quality;
- loss of volume of soil;

---

[26] *Our emphasis*: in most applications LCA is subject to a number of simplifying assumptions in order to make the analysis practically achievable. These simplifications can introduce a degree of subjectivity into the analyses.

- loss of volume of groundwater;
- energy consumption (equivalent CO_2);
- air emissions (including CO_2 equivalent greenhouse gases);
- surface water emissions;
- waste creation;
- use of space.

REC is not designed to be a full sustainability appraisal, for example considering a full range of environmental, social and economic factors, but does provide decision makers with a perspective on the environmental sustainability versus direct costs of different remediation options.

20.3.4.15 Quality of Life Capital Assessment

Quality of life capital assessment is a sustainability appraisal tool for maximising and integrating environmental, economic and social benefits as part of any land use or management decision. The core idea of Quality of life Capital is that the environment, the economy and society provide a range of benefits for human life, and that it is these benefits or services which need to be protected and/or enhanced. Assessment examines these benefits and services systematically, using a series of questions:

- who the services matter to, why, and at what spatial scale;
- how important are they, distinct question from the previous one;
- whether the benefits and services are in short supply;
- what (if anything) could make up for any loss or damage to the service.

Expert judgement and community views both need to be reflected, so quality of life capital uses both public consultation and involvement processes and technical appraisal methods including (for environmental benefits and services) environmental impact assessment, landscape, ecological, archaeological and characterisation studies (Countryside Agency et al. 2001). The output of the process is a matrix of written conclusions, rather than a formal valuation.

20.4 Applied Sustainable Remediation

In practice, remediation planning can take place during project design and during remedy selection. During project design, there is an opportunity for enhanced sustainability compared with decisions made at the level of remedy selection alone. Opportunities for improving sustainability at the level of remedy selection only relate to the impacts of the remediation process options, for example their use of energy and resources, their impacts on soil functionality, their direct and indirect costs, or the possibilities that they might cause nuisance or have wider effects on health. Opportunities at the project design stage allow for the development of synergy between the land use, development and remediation components of the project. These opportunities might include, but are not limited to:

- Avoidance of unnecessary remediation work, as in the basement parking example earlier, or simply by the appropriate placement of less sensitive parts of a development so as to reduce the stringency of Risk Management objectives; see Case Study in Section 20.5.1.
- Designing the land use to facilitate longer term, lower input remediation methods, the so called extensive methods (Nathanail et al. 2007); see Case Studies in Sections 20.5.1. and 20.5.2.
- Linking remediation systems to renewable energy; see Case Study in Section 20.5.5.
- Linking land use and built developments to renewable energy production and use (AEA Technology PLC and r3 environmental technology limited 2004).

A consistent and structured approach to considering sustainability allows these opportunities to be identified and exploited. Considering site sustainability is in some ways analogous to considering site risks. The sustainability of a project will have a *status* or *condition* that reflects the benefits and impacts of the project on the various indicators of sustainability. Some form of *assessment* allows an evaluation of this sustainability. Different *interventions* allow sustainability to be *managed* to perhaps reach particular sustainability objectives or goals for a site, or simply to maximise the sustainability that can be achieved. Hence, sustainable remediation includes stages of *assessing* and *managing* sustainability. This may well be a cyclic process. As remediation work continues and either changes the site circumstances, or leads to new knowledge, project objectives may be revisited. So the consideration of Risk Management and sustainability issues may be an iterative process.

Since at present no general guidance exists in many countries, the approach for sustainability appraisal needs to be agreed on a case by case basis. Supporting frameworks are being developed to fill this gap (such as the developing SURF, NICOLE and SuRF-UK approaches – see Table 20.3), but the choice of which indicators to consider will largely depend on the project, its circumstances and the stakeholders involved. Hence, the first step in sustainability appraisal is to discuss how sustainability will be incorporated along the decision making process and which indicators are suited to describe the desired sustainability performance. A qualitative approach can be used to screen the "sustainability" of a range of remediation options. This may be sufficient to provide a clear cut decision, but if this is not possible, the qualitative step should help to refine the remediation option shortlist and clarify what are the sustainability issues of particular importance for the project.

These issues of importance may be linked to strong corporate or local planning preferences, for example, the protection of local amenity green space, the avoidance of road traffic or particular targets related to the carbon intensity of operations. These are in effect the KPIs (Key Performance Indicators) for a particular project. The selection of KPIs is likely to be strongly context specific and determined by negotiation, affected by the project and its circumstances and the corporate, planning and other drivers affecting the particular set of decision makers involved. While major environmental policies mean that there will be some issues in common across projects, such as climate change, others, such as Biodiversity impacts may be project dependent.

Once the important sustainability issues, *in the view of the stakeholders*, have been agreed, their effect on decision making also can be considered. For example, what is the optimum solution if the remediation costs of a certain technology are 10% higher but the CO_2 emissions are 20% lower than a feasible alternative?

A part of this discussion may include deciding what outcomes are compulsory (for example achieving Risk Management) and what outcomes are expected, but do not form part of the regulatory permit or contractual regime for the remediation project. The range of solutions available and the scope of any improvement in sustainability depend strongly on the stage of project planning. For example, at early stages changes in general project approach may have strong sustainability benefits from the standpoint of remediation. However, whether or not these changes take place is dependent on a trade-off between remediation planning and other aspects of project planning. It may be necessary to quantify different approaches using cost benefit analysis. This analysis may be limited, on grounds of cost and complexity, to a limited set of key sustainability drivers for the project. If many projects decisions are already made and cannot be easily revisited, there is reduced flexibility in achieving sustainable remediation. It needs to be recognized that some decisions will be made in this context. For example, contractors requested to bid for a particular remediation project will not be able to alter the boundaries of that project in their tender submission. However, these contractors could identify technologies that have a lower impact on the environment, e.g. by a reduced need for transportation, energy usage, etc.

An emerging debate is sustainability of remediation work in terms of its *robustness* over time, as environmental conditions change, or to facilitate changing land use requirements in the future (appropriate institutional controls). This idea of robustness is not a re-emergence of the “multi-functionality” debate. It is generally agreed now that it is most prudent to tailor remediation to actual future land use, rather than *any* future land use (Vegter et al. 2002). Rather, robustness considers realistic future scenarios in terms of land use and climate change, for example, and take a view on whether it might be beneficial to consider them during project design and remedy selection. The aim of this consideration is to design with flexibility and at reasonable costs, if it will demonstrably reduce the likelihood of project inadequacy or failure in the medium term. Table 20.6 summarises some hypothetical effects of climate change on contaminant mobility and hence Risk Management needs. These effects might result from higher temperatures, more intensive rainfall over short periods and longer dry periods (CLAIRE 2007).

Table 20.6 Indicative examples of possible impacts of climate change on the mobility of contaminants (not an exhaustive listing)

Process	Impact on contaminant mobility
Increased leaching/erosion	Increase
Increased soil drying	Increase
Increased evaporation	Increase
Change in redox condition	Increase/Decrease
Enhanced biodegradation	Decrease
Affecting mineral capping (bentonite)	Increase
Increased flooding risk	Increase

20.5 Case Studies

The case studies following are either examples of remediation carried out with additional sustainable development "gain" in mind, or examples of the use of sustainability appraisal tools in decision making.

20.5.1 Soil Redevelopment in the Volgermeerpolder, Amsterdam, the Netherlands

In the 19th and 20th century peat was an important fuel in the Netherlands. North of Amsterdam peat layers with a thickness of up to 8 m were excavated. In the Volgermeerpolder a 100 ha wetland remaining after this peat extraction was used as a chemical and domestic waste dump for the municipality of Amsterdam until 1980. Residues from chlorinated pesticide production were also deposited in the Volgermeerpolder, which resulted in contamination of the surrounding waterways with dioxins. Groundwater on the site is severely contaminated with aromatic and chlorinated compounds, mainly benzene and chlorobenzene, both biodegradation products of the pesticides. However, the residual peat in the subsurface strongly sorbs organic solutes and allows limited migration of the contaminants, resulting in a fragile but stable situation without off-site migration of contaminants.

A large scale remediation project for this area was initiated in 2003. From the beginning of the project planning it was clear that a remediation of the site by removal of the waste would not be feasible. A remediation concept was developed centred on the reclamation of the area as a constructed wetland with bog development. The waste deposit will be capped, using a low permeable layer (PE-liner). This cover layer will prevent infiltration of rain water, and reduce the contaminant flux into the surrounding subsurface. The wetland system will be constructed on top of this cover layer (see Fig. 20.9 for an aerial view of the Volgermeerpolder). It consists of 60 individual wetland cells ("paddy fields"), separated by dykes with an integral water management system similar to the construction of rice fields in Asia.

Fig. 20.9 Aerial view of the Volgermeerpolder site[27]

The aim is the gradual development of a bog in the area. A wider sustainability benefit of this bog development is the creation of a 100 ha carbon sink. The area will also be suitable for recreational purposes, and will provide valuable habitat area.

20.5.2 Wind Powered Passive Aeration Remediation Systems

Wind energy can be converted to electrical energy which could power a remediation process. However air itself is a medium that consists of the most elemental parameter to enhance aerobic biodegradation, oxygen. Using the energy from wind to inject air into the soil is a very efficient method to enhance aerobic bioremediation. The energy from wind can be applied to inject air into the unsaturated upper soil (*venting*) with moderate to high hydraulic conductivities. However, the overpressure that can be created is generally insufficient to inject air in the saturated zone (*sparging*).

Wind powered turbine ventilation systems have been applied to the in situ aerobic degradation of petroleum hydrocarbons. For example, at site in the Netherlands, Tauw, found that residual soil contamination with a light petroleum hydrocarbon fraction remained after excavation of the major sources (unpublished information). The residual contamination was present in the unsaturated upper soil layer. The subsurface consisted of moderately coarse sands with a hydraulic conductivity of 3–5 m/day. The vadose zone was aerated by a wind powered turbine system, using equipment generally applied to indoor air ventilation in houses (see Fig. 20.10).

Other passive venting systems that can be used are based on differences in barometric pressures (see Fig. 20.11 for a passive soil venting system tested

---

[27] De Jong Luchtfotografie (2009) www.dejongluchtfotografie.nl

Fig. 20.10 Typical wind-powered ventilation system


A PSVE well with a BaroBall valve

Photos provided by Savannah River Site

Fig. 20.11 Passive soil venting system tested at the Savannah River Site, USA

at the Savannah River Site, USA). These pressure differences can either result in injection or extraction of air from soil, thereby extracting contaminants or enhancing natural degradation processes by the injection of fresh air (Riha 2003).

20.5.3 Sustainable Reuse of Contaminated Sediments

Sediments of the Old IJssel River system in the east of the Netherlands are severely contaminated due to historical industrial activities. The rivers are regularly dredged, and their sediments are contaminated. The dredged sediment is usually stored in a temporary depot for dewatering. After dewatering, which may take a few months to a year, the drained sediments are excavated and transported to a landfill.

These sediments have been re-used by a local authority project to substitute for a primary resource (sand) for a 3,700 m^2 car park at a recreational area, in the province of Utrecht. 1,500 m^3 of dredged sediment were dewatered in a temporary landfill leaving a volume of 1,000 m^3 drained sediment which were then mixed with a proprietary stabilisation amendment Megatrax®. The pavement met the leaching criteria of the Dutch Building Materials Decree

The re-use of the sediment avoided the use of 1,000 m^3 of a primary resource and resulted in a CO_2 emission reduction of 15 % compared with the use of primary raw materials.

20.5.4 The Use of the REC Method to Select a Remediation Strategy

The REC (Risks, Environmental Merits and Costs) method was used in the remediation strategy development for a large chemical production plant in the province of Zeeland in the Netherlands, where historical incidents have led to multiple soil contamination hotspots, including aromatic hydrocarbons and chlorinated solvents. These hotspots have resulted in several groundwater contaminant plumes, which are partly intermingled.

The following remediation alternatives were compared, the using REC method:

- complete removal of contaminations, i.e. a multi-functional approach (referred to as “complete removal alternative”);
- monitoring Natural Attenuation in combination with local hot spot removal (referred to as “buffer zone alternative”)
- hydrological containment including removal of mobile light non-aqueous phase liquid (LNAPL) contaminants (referred to as “IBC+ alternative”);
- hydrological containment (referred to as “IBC-alternative”).

The buffer zone alternative is a risk-based remediation approach and consists of monitored Natural Attenuation in combination with local hotspot removal. The buffer zone is defined as the zone at the site boundary in which Natural Attenuation is able to reduce the contaminant concentrations at the site boundary to the values accepted by the local authorities (Dutch T-value) so that the potential off site migration of the contamination does not cause any risks outside the site. The hydrological containment strategy would involve pumping groundwater at a rate of 25 m^3/h with 18 deep wells on site.

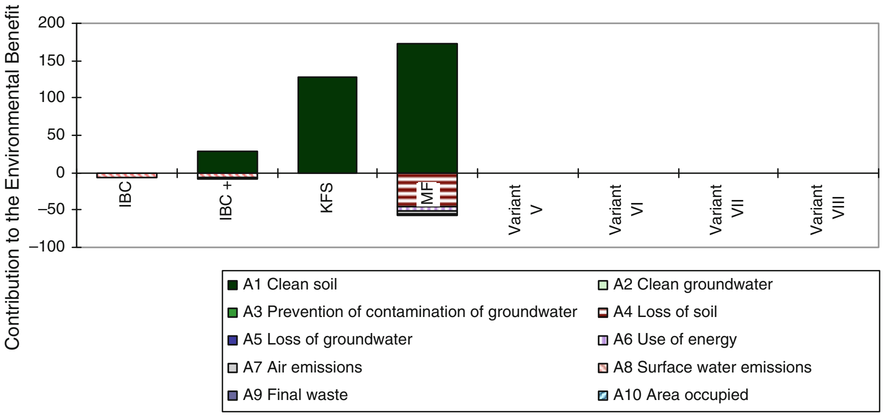


Fig. 20.12 An example comparison of the environmental merits, both positive and negative aspects for the four remediation alternatives (IBC = hydrological containment; KFS = buffer zone alternative; MF = complete removal alternative)

The REC method was used to quantitatively compare the environmental effectiveness of the four remediation alternatives. The environmental merit index encompasses positive aspects (such as clean soil and groundwater) and negative aspects (such as energy consumption, use of clean water and space, contamination of other environmental compartments). The environmental merit is based on the principle that the negative impact on the environment has to be as limited as possible and that the supplies of raw materials have to be preserved for future generations to the greatest possible extent. Figures 20.12 and 20.13 show the environmental merit for the four remediation options. Figure 20.14 compares environmental merit with the costs of the four remediation alternatives.

Environmental merit was highest for the buffer zone alternative and the complete removal alternative, with the buffer zone alternative showing a slightly more positive

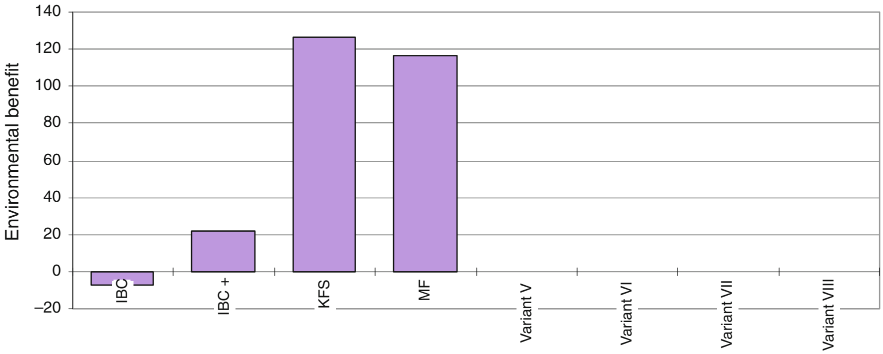


Fig. 20.13 Comparison of the total environmental merit of the remediation alternatives (IBC = hydrological containment; KFS = buffer zone alternative; MF = complete removal alternative)

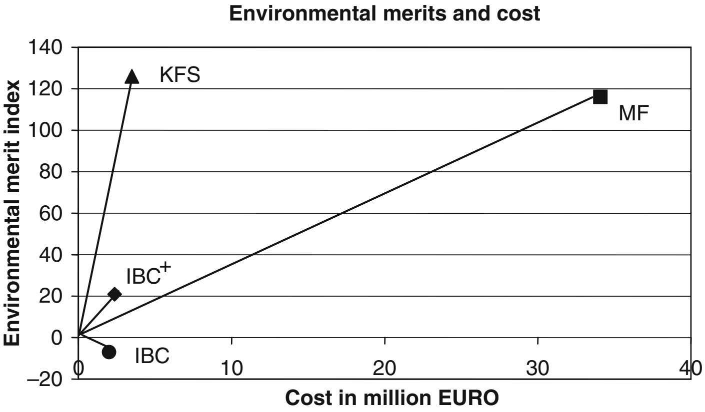


Fig. 20.14 Comparison of remediation alternatives (IBC = hydrological containment; KFS = buffer zone alternative; MF = complete removal alternative)

environmental benefit, than the complete removal alternative (Fig. 20.13). As illustrated in Fig. 20.12 the explanation is that, although the buffer zone alternative has a lower value for the aspect "clean soil" than the complete removal alternative, aspects with negative environmental values show only very low values compared to those for the complete removal alternative.

The IBC+ alternative (hydrological containment) strategy has a negative overall environmental benefit, but was the least expensive and has a negative environmental benefit, and leaves a long term liability from the contamination remaining *in situ*. The estimated costs of complete removal of contaminations (complete removal alternative) were considerably higher compared to a buffer zone alternative with a remediation up to the T-value. As can be seen from the slope of the alternatives a buffer zone alternative remediation up to the T-value will have the highest environmental effectiveness (cost versus environmental benefit). Both the site-owner and the local authorities expressed a preference for the buffer zone alternative, since it involved active remedial measures.

20.5.5 "Sanergy" as a Sustainable Synergy of Remediation and Groundwater Energy

A former industrial site in Eindhoven, the Netherlands, is being redeveloped into a business and housing area. Buildings were planned to be heated and cooled with groundwater linked heat pumps to provide sustainable energy and reduce their CO_2 emissions. However, the groundwater is contaminated as a result of decades of industrial activity. At an early stage of project design, the idea was conceived to use the groundwater flows of the energy system also for the containment and remediation of the contamination.

Energy can be obtained from groundwater by abstracting groundwater and utilizing it's thermal value (heat or cold) with a heat pump. The conventional approach for Heat-Cold-Storage (HCS) is that groundwater is pumped from a cold zone to a warm zone. The combination with remediation demanded a change in this approach, because contaminants would have been moved and spread as well. Conventionally groundwater remediation systems are designed to contain and reduce the extent of contaminants.

Synergy of groundwater energy and remediation is possible if the necessary large flows are used to contain the contaminants. Instead of using cold and warm zones in the subsurface it was decided to use a so called recirculation system. This system uses a continuous flow direction and extracts heat or cold from groundwater with a constant temperature (see Fig. 20.15). The remediation effect is a result of the enhancement of naturally occurring degradation. Naturally occurring bacteria, electron donors, nutrients and contaminants are more effectively mixed by the large groundwater flows. The development of degradation conditions will be intensively monitored, and adjusted if necessary with the injection of electron donor.

This combined approach leads to reductions in CO_2 emissions and use of non-renewable energy of 30–50% compared with traditional heating. Furthermore, the containment and remediation of the groundwater contamination does not require additional energy, and is accomplished by the groundwater energy system. The remediation is sustainable in a sense that it does not place any additional burden on the environment and there is no greenhouse gas production or consumption of non renewable sources.

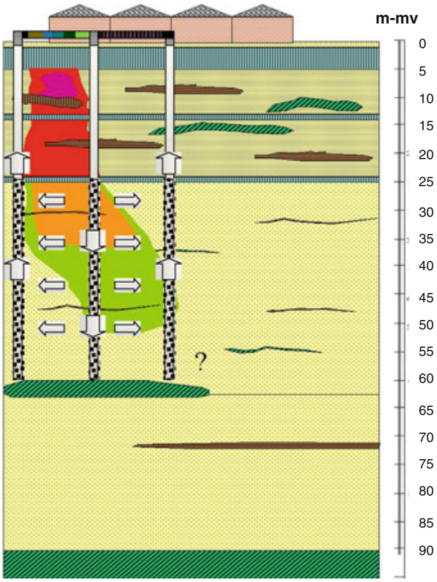


Fig. 20.15 "Sanergy" schematic diagram: extraction wells surrounding the infiltration in the middle of the contaminant plume

20.6 The Future Perspective of Sustainable Management of Contaminated Sites

20.6.1 A New Basis for Decision Making

Over the last 10 years or so contaminated land management approaches have moved towards a risk based land management procedure. Currently, another threshold must be crossed towards more sustainable land management, which encompasses Risk Management as land management can only be sustainable if it manages at least to remove unacceptable risks. Protection of human health, water and the wider environment remain the over-riding priority, but the negative impacts of over engineering responses also needs to be taken into account. In practical terms, risks from soil contamination may be low, controllable or negligible for key receptors. In this situation a more holistic view of the need for removal or treatment of contamination is important, for example taking into account the resource and energy intensity of remediation. In some cases this might imply that we will move from "intensive" soil remediation to monitoring and management, simply because it is the best option for the overall environment. The biggest leap forward is not the development of new techniques, but the development of better tools for option and sustainability appraisal, increasing awareness of sustainability as an important decision making issue and communicating these new concepts. This will require an impetus at national and international levels, so that site owners, service providers and regional and local authorities and regulators can confidently follow with sustainable solutions. The first steps in this international sustainable remediation debate are already underway.

20.6.2 Work in Progress

"Sustainable remediation" is an emerging concept. Sustainability will often lie more in the evaluation of options for Risk Management, rather than promoting any particular technological approaches. The full potential for sustainability in remediation may be difficult to achieve if remediation choices are regarded as being at the end of a pipeline of site use and project planning decisions that have already been taken. It is important for the sustainability of remediation requirements to take place *early* during site and project management consideration. A prerequisite for the sustainable remediation debate is a common understanding of terms, for example agreeing on a scope of what is to be considered. The current international debate appears to be reaching a consensus in terms of principles, but has not reached a consensus in terms of a common understanding of the scope of sustainable remediation or the terms that are used in it. Another debate that is just beginning is the relationship between the management of "sustainability" and "risks". This is a complicated debate that will take time to resolve. For example, lower intensity remediation approaches may not lead to the mass removals expected by some regulators, or may not meet particular statutory goals. However, these methods may still achieve practical Risk Management. Perhaps a particular generic limit value may be overly conservative for a particular site. Indeed, in some cases it may well be argued that a mass removal-based remedial objective is a poor result for the environment, because the energy and diesel consumption, green house gas emissions and traffic risks may be far worse than the presence of immobile contaminant in the subsoil. What is needed is a more flexible and adaptive approach to remediation and Risk Management that is focussed on the sustainable reduction of risks rather than the achievement of particular threshold values.

20.6.3 Technological Innovation by Combining State of the Art Techniques

There may be major technological advances that change the way remediation is carried out in the future. However, in the short term is also a large potential to improve sustainability using existing technologies. A wide range of sustainable remediation opportunities can be created by applying and combining existing state of the art techniques that are not necessarily from the field of remediation. In particular, important steps forward will be possible through intelligent design and synergy, as illustrated in the case studies, for example:

- technologies, like solar or wind energy for energy consuming remediation techniques;
- groundwater energy systems and groundwater remediation;
- multiple subsurface use, subsurface building and contaminant removal or contaminant immobilisation or contaminant isolation;
- nature restoration, management and policy development.

The applicability and benefits are generally site-specific. In most cases remediation will put a burden on the environment, for example relating to the inputs needed to install the system. These burdens need to be balanced against the benefits delivered by the remediation.

20.6.4 Synergies: Go with the Flow

The importance of the climate change and sustainability debate may be an opportunity for enabling renewed interest and investment in remediation where useful synergies between remediation and other forms of sustainable development such as renewable energy can be exploited. An emerging area of interest across Europe and the USA is the use of Brownfield sites for renewable energy production, for example thorough the cultivation of biomass (Bardos et al. 2009). Of particular interest in the Netherlands is the synergy between groundwater energy management and groundwater remediation.

In many urban areas there is extensive groundwater contamination from a wide variety of sources that result from decades if not centuries of activities. Remediation of this groundwater resource alone is not seen as technically or financially feasible. However, the use of groundwater heat storage will pump huge amounts of groundwater. Rather than seeing groundwater protection legislation as a hindrance to sustainable energy, perhaps sustainable energy should be seen as an opportunity for groundwater improvement for little additional use of energy or resources. This concept is already under active discussion in the Netherlands. In many Dutch cities the benefits of energy and emission reduction are thought to be more important than the mixing of contaminated groundwater. In the city of Utrecht in the Netherlands a huge groundwater management zone is planned for this purpose. Schemes are also under consideration in the Dutch cities of Apeldoorn, Tilburg and Zwolle. The functional use of groundwater is therefore going to lead to rapid restoration of groundwater quality. As a result of increased dynamics and mixing, the groundwater quality is also expected to improve. The former premises of Philips in Eindhoven (Case Study in Section 20.5.5) will be the first site in the world to have such a dual system operating.

References

AEA Technology PLC and r3 environmental technology limited (2004) Markham Willows masterplanning, exSite Research Ltd, Hillam, Leads, UK. http://www.r3environmental.co.uk/downloadsnew/index.asp. Accessed January–March 2009

Bardos RP (2003) Report of the NICOLE workshop: management of contaminated land towards a sustainable future: opportunities, challenges and barriers for the sustainable management of contaminated land in Europe, 12–14 March 2003, Barcelona, Spain, Contamin Reclam 11(4):449–472

Bardos RP (2008a) Report of the NICOLE/SAGTA workshop: sustainable remediation 3rd March 2008, London, UK. Land Contamin Reclam 16(4):381–403

Bardos RP (2008b) Report of the NICOLE workshop: environmental decision support systems 9–10 October 2008 Madrid, Spain, NICOLE Secretariat, Deltares/TNO, Apeldoorn, the Netherlands. http://www.nicole.org/publications/library.asp?listing=1. Accessed January–March 2009

Bardos P, Chapman A, Andersson-Sköld Y, Blom S, Keuning S, Polland M, Track T (2009) Compost on marginal land for biomass production. Biocycle 49(12):50–52

Bardos RP, Lewis AJ, Nortcliff S, Mariotti C, Marot F, Sullivan T (2002) Review of decision support tools for contaminated land management, and their use in Europe, Final report. Austrian Federal Environment Agency, 2002 on behalf of CLARINET, Spittelauer Lände 5, A-1090 Wien, Austria. Available from http://www.umweltbundesamt.at/en/umweltschutz/altlasten/projekte1/international1/clarinet/clarinet_results/. Accessed January–March 2009

Bender T, Volkwein S, Battermann G, Hurtig HW, Klopffer W, Kohler W (1998) Life cycle assessment for remedial action techniques: methodology and Application. Contaminated Soil’ 98:367–376, Proc 6th Intern FZK/TNO Conference on Contaminated Soil, 17–21 May 1998, Edinburgh, UK. ISBN 07277 26757

Bennett R (2006) Builders swallow up gardens for flats. Ministers are accused of deceiving the public over the use of Brownfield sites for new housing, The Times May 25, 2006. http://www.timesonline.co.uk/tol/news/uk/article725599.ece

Brundtland GH (1987) Our common future, World Commission on Environment and Development, Oxford University. ISBN 0-19-282080-X. http://www4.oup.co.uk/isbn/0-19-282080-X. Accessed January–March 2009

Building Research Establishment (2004) Assessment of sustainability tools, Report number 15961, July 2004, SUE MoT Subcontract: BRE Report, Building Research Establishment (BRE), Glasgow, UK. http://sue-mot.org/publications/

Carbon Trust, Defra and BSI British Standards (2008a) Guide to PAS 2050 How to assess the carbon footprint of goods and services, BSI, 389 Chiswick High Road, London W4 4AL, UK. ISBN 978-0-580-64636-2. www.bsigroup.com/en/Standards-and-Publications/Industry-Sectors/Energy/PAS-2050/

Carbon Trust, Defra and BSI British Standards (2008b) PAS 2050:2008 Specification for the assessment of the life cycle greenhouse gas emissions of goods and services, BSI, 389 Chiswick High Road, London W4 4AL, UK. IISBN 978 0 580 50978 0. www.bsigroup.com/en/Standards-and-Publications/Industry-Sectors/Energy/PAS-2050/

Chambers N, Simmons C, Wackernagel M (2007) Sharing nature's interest: ecological footprints as an indicator of sustainability. Earthscan, London, ISBN 978-1-85383-739-5. http://www.earthscan.co.uk. Accessed January–March 2009

CL:AIRE (2007) Climate change, pollutant linkage and Brownfield regeneration subr:im bulletin SUB3 Contaminated Land Applications in Real Environments (CL:AIRE), London, UK. www.claire.co.uk

CL:AIRE (2009) A review of published sustainability indicator sets: how applicable are they to contaminated land remediation indicator-set development? March 2007. 9 May 2009. Contaminated Land Applications in Real Environments (CL:AIRE), London, UK, ISBN 978-1-905046-18-8, www.claire.co.uk/surfuk

CL:AIRE (2010) A framework for assessing the sustainability of soil and groundwater remediation. Public Consultation March 2010 CL:AIRE, London, UK. ISBN 978-1-905046-19-5

Countryside Agency, English Heritage, English Nature and Environment Agency (2001) Quality of life capital managing environmental, social and economic benefits, Overview report. Natural England, Northminster House, Peterborough, PE1 1UA, UK, http://www.countryside.gov.uk/LAR/archive/Quality/Contacts_reports/reports.asp. Accessed January–March 2009

Dalemo M, Oostra H (1997) The ORWARE Model: Results from a Case Study in Sweden. In: Stentiford EI (ed) Proceeding of the international conference on organic recovery and biological treatment into the next millenium, Harrogate-UK, pp 263–269. ISBN 0952227037

Danish Topic Centre on Waste and Resources (2006) A quick guide to LCA and CBA in waste management, Danish topic centre on waste and resources, Copenhagen, Denmark http://wasteandresources.dk/d/p/LCA-CBA%20eng.pdf. Accessed January–March 2009

Department for Environment Food and Rural Affairs (2005a) Securing the future, delivering UK sustainable development strategy. The UK Government sustainable development strategy Cm 6467 ISBN 0-10-164672-0, The Stationary Office, Norwich http://www.tsoshop.co.uk/. Accessed January–March 2009

Department for Environment Food and Rural Affairs (2005b) One future – different paths, The UK's shared framework for sustainable development. Product code PB10591, Department for Environment Food and Rural Affairs (Defra) Publications, Admail 6000, London, SW1A 2XX http://www.sustainable-development.gov.uk/publications/pdf/SD%20Framework.pdf. Accessed January–March 2009

Department for Environment Food and Rural Affairs (2006b) Carbon balances and energy impacts of the management of UK wastes, Defra R&D Project WRT 237 Final Report. ERM and Golder associates and department for environment food and rural affairs (Defra), London, UK. http://www2.defra.gov.uk/research/project_data/More.asp?I=WR0602&M=KWS&V=Carbon+balance&SUBMIT1=Search&SCOPE=0. Accessed January–March 2009

Department for Environment Food and Rural Affairs (2007b) Consultation on the secretary of state's guidance: general guidance manual on policy and procedures for A2 and B installations, 12 June 2007, Department for environment food and rural affairs (Defra), London, UK. http://www.defra.gov.uk/corporate/consult/envpermitprog4/index.htm

Department for Environment Food and Rural Affairs Defra (2007c) A review of recent developments in, and the practical use of, ecological footprinting methodologies, A research report completed for the Department for Environment, Food and Rural Affairs (Defra) by Risk & Policy Analysts Ltd, Defra, London. http://www.defra.gov.uk/science/project_data/DocumentLibrary/EV02024/EV02024_5880_FRP.pdf

Department for the Environment Transport and the Regions (1998) Planning for the communities of the future (Cm. 3885), ISBN 0 10 138852 7, www.communities.gov.uk/archived/publications/planningandbuilding/planning. Accessed January–March 2009

Department of the Environment, Transport and the Regions (1999) Review of technical guidance on environmental appraisal. A report by EFTEC (Economics for the Environment Consultancy), Defra, London, UK. http://www.defra.gov.uk/environment/economics/rtgea/index.htm. Accessed January–March 2009

Department of the Environment, Transport and the Regions, Environment Agency, Institute for Environment and Health (2000) Guidelines for environmental risk assessment and management. Defra, London, UK. http://www.defra.gov.uk/environment/risk/eramguide/index.htm. Accessed January–March 2009

Edwards D, Bertram C, Pahlen G, Nathanail P (2005) Best practice guidance for sustainable Brownfield regeneration. Land Quality, a Division of Land Quality Management Ltd, Nottingham, UK. ISBN 0-9547474-0-2

Environment Agency (1999) Cost benefit analysis for remediation of land contamination, R&D technical report P316, prepared by Risk Policy Analysts Ltd and WS Atkins. Available from: Environment Agency, Bristol, UK. ISBN 185705 0371

Environment Agency (2000) Assessing the wider environmental value of remediating land contamination. Environment Agency R&D Technical Report P238, available from: Environment Agency, Bristol, UK. ISBN 185705 0371 http://www.eugris.info/displayresource.asp?ResourceID=3653&Cat=document. Accessed January–March 2009

Environment Agency (2000b) Costs and benefits associated with remediation of contaminated groundwater: framework for assessment. R&D Technical Report P279, prepared by Komex Clarke Bond & EFTEC Ltd., available from: Environment Agency, Bristol, UK

Environment Agency (2009) Waste and resources assessment tool for the environment (WRATE). http://www.environment-agency.gov.uk/research/commercial/102922.aspx. Accessed 17 Apr 2009

EURODEMO Consortium (2007a) Environmental efficiency criteria. Report on case studies, Deliverable reference number: D 5–2, Project no. (GOCE) 003985, Austrian Environmental Protection Agency, Vienna. http://www.eurodemo.info/results/. Accessed January–March 2009

EURODEMO Consortium (2007b) Framework for sustainable land remediation and management. Deliverable reference number: D 5–3, Project no. (GOCE) 003985, Austrian Environmental Protection Agency, Vienna. http://www.eurodemo.info/results/. Accessed January–March 2009

Finnamore JR, Denton B, Nathanail CP (2000) Land contamination – management of financial risk, C545, CIRIA, London, UK, ISBN 979 0 86017 5452

Finnveden G (2000) On the limitations of life cycle assessment and environmental systems analysis tools in general. Int J Life Cycle Assess 5(4):229–238

Fischer TN (2007) Theory and practice of strategic environmental assessment. Earthscan, London, UK. ISBN 978 1 84407 4525

Global Reporting Initiative (2006) Sustainability reporting guidelines, Version 3, Global Reporting Initiative, 1001 EA Amsterdam, The Netherlands. www.globalreporting.org

Hanley N, Spash C (1994) Cost benefit analysis and the environment. Edward Elgar, Aldershot, Hampshire. ISBN 1-85278-947-6

Harbottle MJ, Al-Tabbaa A, Evans CW (2008a) Sustainability of land remediation, Part 1: overall analysis, Proceedings of the institution of civil engineers geotechnical engineering 161, April 2008 Issue GE2: 75–92, http://www.atypon-link.com/doi/pdf/10.1680/geng.2008.161.2.75

Harbottle MJ, Al-Tabbaa A, Evans CW (2008b) Sustainability of land remediation, Part 2: impact assessment, Proceedings of the institution of civil engineers geotechnical engineering 161, June 2008 Issue GE3: 117–127, http://www.atypon-link.com/doi/pdf/10.1680/geng.2008.161.3.117

Hazebrouck B, Ledrans M (2005) Stakeholder participation in site assessment and management: a real case of Vincennes, France, Consoil 2005, Proceedings of the 9th international FZK/TNO conference on Soil Water Systems. 3–7 October 2005 Bordeaux: 29

HM Treasury The green book appraisal and evaluation in central government treasury guidance, The stationery office, London. http://www.hm-treasury.gov.uk/d/2(4).pdf. Accessed January–March 2009

Koneczny K, Dragusanu V, Bersani R, Pennington DW (2007) Environmental assessment of municipal waste management scenarios: part I – data collection and preliminary assessments for life cycle thinking pilot studies. European commission joint research centre, Institute for environment and sustainability, JRC Ispra, Italy, JRC 41238, Report EUR 23021 EN, ISBN 978-92-79-07449-3, Office for Official Publications of the European Communities, Luxembourg

Koneczny K, Pennington DW (2007) Environmental assessment of municipal waste management scenarios: part II – detailed life cycle assessments. European commission joint research centre, Institute for environment and sustainability, JRC Ispra, Italy, JRC 41241, Report 23021 EN/2, ISBN 978-92-79-07450-9. Office for Official Publications of the European Communities, Luxembourg

Lowe EA, Warren JL, Moran SR (1997) (eds) Discovering industrial ecology. An executive briefing and sourcebook, Battelle, Columbus, OH. ISBN 1574770349

Ministry of Defence (2006) MOD sustainability and environmental appraisal tools handbook. Defence estates, Version 4.0, December 2006, Room 14, K Block, Foxhill, Bath, BA1 5AB. http://www.mod.uk/DefenceInternet/AboutDefence/CorporatePublications/DefenceEstateandEnvironmentPublications/DefenceEstates/SustainabiltyAndEnvironmentalAppraisalToolHandbook.htm

Mitchell G (2005) Assessing the sustainability of development options for outer urban areas: some issues for the SOLUTIONS project, EPSRC SOLUTIONS project 2nd annual conference, UCL, December 2005, http://www.suburbansolutions.ac.uk/DocumentManager/secure0/07B_Assessing%20the%20sustainability-Some%20Issues%20(paper).pdf

Mulberg J (1996) Environmental planning, economic planning and political economy. Int Sociol 11(4):441–456

Nathanail CP (2009) Sustainable remediation: from promise through policy to practice. Presentation to international committee on contaminated land. 16 and 17th of September, 2009 Helsinki, Finland

Nathanail CP, Bardos RP (2004) Reclamation of contaminated land. Wiley. ISBN 0471985600 http://www.wileyeurope.com/WileyCDA/WileyTitle/productCd-0471985619.html. Accessed January–March 2009

Nathanail J, Bardos P, Nathanail P (2007) Contaminated land management ready reference, Update EPP publications/Land Quality. Available from: EPP Publications, 6 Eastbourne Road, Chiswick, London, W4 3 EB. E-mail enquiries@epppublications.com. ISBN 1900995069. www.readyreference.co.uk

NICOLE (1999) Communication on contaminated land. Network for industrially contaminated land in Europe, http://www.nicole.org/. Accessed January–March 2009

Nichols RL, Looney BB (2007) First, do no harm. Proceedings of the 2007 Georgia water resources conference, held March 27–29, 2007, at the University of Georgia. WSRC-STI-2006–00352, http://cms.ce.gatech.edu/gwri/uploads/proceedings/2007/87.pdf

NOBIS (1995a) Risk reduction, environmental merit and costs, REC-method, phase 1, Dutch research program biotechnological in situ remediation, Document 95-1-03, CUR/NOBIS, Gouda, the Netherlands

NOBIS (1995b) Risk reduction, environmental merit and costs. REC-method, phase 2: a methodology based on risk reduction, Environmental merit and Costs. Dutch research program biotechnological in situ remediation, Document 95-1-03, CUR/NOBIS, Gouda, the Netherlands

Office of the Deputy Prime Minister (2004) Planning policy statement 23: planning and pollution control. Annex 2: Development on land affected by contamination, HMSO, St Clements House, 2–16 Colegate, Norwich NR3 1BQ, UK. ISBN 0 11 753932 5 (Now "owned" by the Department for Communities and Local Government – DCLG) www.communities.gov.uk/publications/planningandbuilding/pps23annex2. Accessed January–March 2009

Office of the Deputy Prime Minister (2005a) Sustainability appraisal of regional spatial strategies and local development frameworks, ODPM, London, UK. ISBN 1851127984 http://www.communities.gov.uk/index.asp?id=1161341. Accessed January–March 2009

Office of the Deputy Prime Minister (2005b) A practical guide to the strategic environmental assessment directive, September 2005, Reference Number 05 PD 03311, ISBN 1851127887, ODPM, London, UK. http://www.communities.gov.uk/publications/planningandbuilding/practicalguide

Orr (2001) Scoping study on public perception of risk and EP risk communication, R&D Technical Report P6–019/2/TR, download at http://publications.environment-agency.gov.uk/epages/eapublications.storefront

Petts J, Homan J, Pollard S (2003) Participatory risk assessment: involving lay audiences in environmental decisions on risk literature review and stakeholder interviews, R&D technical report E2-043/TR/01, download at http://publications.environment-agency.gov.uk/epages/eapublications.storefront

Riha BD (2003) Passive soil vapor extraction (PSVE) for VOC remediation at the metallurgical laboratory (MetLab) June 2005 progress report, WSRC-TR-2005–00268 Prepared for the US Department of Energy under Contract DE-AC09–96SR18500 http://www.durhamgeo.com/pdf/Rem-pdf/Other/WSRC-TR-2005–00268%20MetLab.pdf

Shakweer A, Nathanail CP (2003) LCA methodology for remediation strategy selection. ConSoil 2003 8th International FZK/TNO Conference

Sustainable Remediation Forum – SURF (2009) Integrating sustainable principles, practices and metrics into remediation projects. Remed J 19(3):5–114. doi:10.1002/rem.20210

Therivel R (2004) Sustainable urban environment – metrics, models and toolkits: analysis of sustainability/social tools. Report to the sue-MoT consortium. 9 June 2004 http://sue-mot.org/publications/. Accessed January 2009

US Environmental Protection Agency (2006) Life cycle assessment: principles and practice. Report EPA/600/R-06/060 May 2006 http://www.epa.gov/nrmrl/lcaccess/pdfs/600r06060.pdf

US Environmental Protection Agency (2008) Green remediation: incorporating sustainable environmental practices into remediation of contaminated sites. Technology Primer April 2008 Report EPA 542-R-08-002. http://www.cluin.org/download/remed/Green-Remediation-Primer.pdf

Vegter J, Lowe J, Kasamas H (eds) (2002) Sustainable management of contaminated land: an overview. Report Austrian federal environment agency, 2002 on behalf of CLARINET, Spittelauer Lände 5, A-1090 Wien, Austria. Available from: http://www.commonforum.eu/publications_clarinet.asp

Waste and Resources Action Programme (2007) The net waste method – testing a new standard for measuring waste neutrality. The Old Academy, 21 Horse Fair, Banbury, Oxon, OX16 0AH, UK. www.wrap.org.uk/construction

Water UK (2008) Sustainability Indicators 2007/08, May 2008, Water UK, Water UK 1 Queen Anne's Gate, London SW1H 9BT. www.water.org.uk

Water UK (2008) Summary of the process to review the Water UK sustainability indicators. Review of the Water UK sustainability indicators May 2008. Water UK. Water UK 1 Queen Anne's Gate, London SW1H 9BT. www.water.org.uk

Waterwise (2007) Hidden waters. A briefing February 2007, Waterwise 1 Queen Anne's Gate London SW1H 9BT, UK. http://www.waterwise.org.uk/images/site/EmbeddedWater/hidden%20waters%2C%20waterwise%2C%20february%202007.pdf

Wittmaier M, Langera S, Sawillaa B (2009) Possibilities and limitations of life cycle assessment (LCA) in the development of waste utilization systems – Applied examples for a region in Northern Germany. Waste Manage 29(5):1732–1738

Wrisberg N, Udo de Haes HA, Triebswetter U, Eder P (eds) (2002) Analytical tools for environmental design and management in a systems perspective. Report of the CHAINET Project (European Network on Chain Analysis for Environmental Decision Support) to the European Union, October 2000. Springer Publications. ISBN: 978-1-4020-0453-7, http://www.leidenuniv.nl/cml/ssp/projects/chainet/#download

Chapter 21
In Situ Remediation Technologies

Tim J.T.C. Grotenhuis and Huub H.H.M. Rijnaarts

Abstract A summary of two decades of developments of In Situ remediation is presented in this chapter. The basic principles of In Situ technology application are addressed, such as equilibrium relations between contaminant phases, factors controlling biological and geochemical processes, contaminant characteristics affecting reductive and oxidative conversion parameters and chemical and biological availability. A wide range of In Situ technologies are discussed within the framework of Risk Management. Technologies can be oriented at contaminant sources, migration pathways or at the receptors. Integration of In Situ technologies in sustainable Risk Management approaches is further evaluated relating to the latest concepts and frameworks. Examples are given of application of In Situ technologies in Risk Management approaches, including those for large scale contaminated Megasites. In the future, one can foresee the rise of combined sustainable technologies for soil, groundwater, surface water and energy.

Contents

T.J.T.C. Grotenhuis (✉)
Wageningen University, Wageningen, The Netherlands
e-mail: tim.grotenhuis@wur.nl

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_21,

21.1 Introduction

21.1.1 Background of In Situ Remediation

In the early 1990s, *In Situ remediation* of contaminated sites started to develop as it became clear that many sites are too large to excavate and treat thereafter. Especially In Situ bioremediation attracted attention after the first publications by McCarty on the sequential biodegradation of TCE (trichloroethylene) at the St. Joseph site near lake Michigan (McCarty et al. 1991) and the prosperous outlook for bioremediation shortly thereafter (McCarty 1991). Besides bioremediation, physical technologies were also developed in that period, such as *pump and treat*, *soil vapor extraction* and *air sparging* (Hutzler et al. 1991; Mackay and Cherry 1989; Marley et al. 1992). Subsequently, bioremediation and physical technologies were combined, such as pump and treat technologies which were used to introduce nutrients or co-substrates to stimulate bioremediation, and soil vapour extraction and air sparging were combined with bioremediation into bioventing technologies (Hoeppel et al. 1991; Malina et al. 1998).

Also more energy intensive technologies were proposed, such as *electro reclamation* (Lageman et al. 1989), and *thermal treatment* and *vitrification* (Dragun 1991). However most In Situ applications in the 1990s were in the field of bioremediation combined with physical removal strategies.

Since the early 1990s, when the first In Situ field studies were performed, it became clear that In Situ processes are rather slow and residual concentrations were often above legal Remediation objectives. Together with the rapidly increasing number of sites detected as contaminated, a transformation in policies and in field applications was observed from removal of all contaminants to reduction of risk at affordable costs. The *bioavailability* of contaminants is a key factor in Risk Based In Situ bioremediation. From the risk perspective, bioavailability determines whether there is a risk for specific receptors. From the remediation perspective the availability of contaminants determines whether in-situ remediation, and to what residual concentration, is possible (Cuypers et al. 2000; Semple et al. 2003; Volkering et al. 1998).

In research, technologies for the concept of Source-Path-Receptor of contaminants in soil and groundwater started to develop. Especially in the period 1995–2005 most attention was paid to mobile contaminants like BTEX (benzene, toluene, etheylbenzene, xylene) and chlorinated solvents. Many concepts in blocking the path of the flow of contaminants within the groundwater were developed, such as bioscreens, biobarriers, funnel and gate systems, permeable reactive barriers.

Technologies for removal of hydrophobic organic contaminants, such as PAHs (polycyclic aromatic carbons), PCBs (Polychlorobiphenyls) and chlorinated pesticides, received less attention in this period. However, this development was also the result of many availability studies in which it was shown that many hydrophobic organic contaminants are often tightly bound to soil organic matter. When such contaminants are tightly bound to soil, the risk for uptake by a receptor is assumed to be minimal. Ironically, an increasing number of research papers show that, if the risk for uptake in the food chain is minimal, also the chance for removal of contaminants is minimal. This implies that high residual concentrations, when proven to be non-available, should not lead to a high risk for human beings, soil organisms, groundwater or agricultural crops. Such a strategy was also studied for heavy metal immobilization at field scale. Different soil additives with a high binding capacity for heavy metals had a positive impact on re-vegetation in highly heavy metal contaminated sites (Vangronsveld et al. 1996).

In 1997, *Natural Attenuation* (NA) caused a breakthrough at the "International Symposium In Situ and On Site Bioremediation", in New Orleans. A development of Natural Attenuation to *Monitored Natural Attenuation* (MNA) and *Enhanced Natural Attenuation* or Engineered Natural Attenuation (ENA) was observed in the subsequent symposia in the USA as well as in Europe at the ConSoil conferences.

Technologies for In Situ source treatment by chemical oxidation of soil contaminants were initially introduced at the laboratory scale, such as (Gates and Siegrist 1995) and (MacKinnon and Thomson 2002) who performed lab scale experiments with MnO_4^- for oxidation of PCE (Perchloroethylene) pools of pure product. It was stated by ISTRC (http://www.estcp.org/) that up to 1999 no field designs were available for In Situ oxidation. Also the use of Fenton's reagent was known for many years as oxidant in waste and wastewater treatment, and was reported to activate the per sulphate oxidation for removal of TCE at lab scale, in 2004 (Liang et al. 2004a,b). Since about 2002, field scale applications of source treatment by *In Situ Chemical Oxidation* (ISCO) were performed with success (Plaisier et al. 2003), leading to a review of the use of Fenton reagent for soil remediation (Pignatello et al. 2006).

21.1.2 Scope

In this Section, first the principles of In Situ remediation will be discussed with respect to equilibrium and non-equilibrium conditions. The primary focus is on In Situ remediation of organic contaminants in the subsoil of the vadose and water saturated zone. No attention is paid to phytoremediation or to In Situ treatment of sites contaminated with heavy metals, nitrates, cyanides or ammonia.

Non-equilibrium conditions are required for remediation, which can be reached by changing specific environmental factors. Often, environmental factors may limit the remediation at a specific site. Therefore, environmental factors are key to the success of an In Situ remediation. The role and possible engineering of several environmental factors will be discussed. Thereafter a variety of Risk Based In Situ technologies will be discussed that can be applied at the source, along the path or at the receptor of the contamination. Finally, an outlook is presented on the combination of In Situ remediation with other societal processes in industrial and urban areas.

21.2 In Situ Remediation Technologies

21.2.1 Principles

21.2.1.1 Equilibrium Relations of Organic Contaminants in Soil

At contaminated sites the contaminants are most often present for a long time before In Situ remediation will be started. Under such conditions, it is often assumed that the contaminants in the pore water are in equilibrium with the soil matrix. Under these conditions specific equilibrium relations are valid.

In many cases, especially when the density of the contaminants is below the density of water, the source of the soil contamination is present in the unsaturated zone (i.e., the zone above the groundwater table). If downward transport occurs, low density contaminants are often present in the water-saturated zone. Especially in the source area of soil contamination, pure product can be present in a pure organic liquid phase (NAPL= Non Aqueous Phase Liquid). In case the density is below the density of water, such as for oils or Total Petroleum Hydrocarbons (TPH), high concentrations of pure product can especially be found floating at the groundwater as LNAPL (Light Non Aqueous Phase Liquid). When the density of the pure product is above the density of water, as in chlorinated solvents, a complex pattern of pure product can be found at large depths upon impermeable soil layers or clay lenses, known as DNAPL (Dense Non Aqueous Phase Liquid) contamination. The NAPL contamination may function as a secondary source of contamination (Norris et al. 1994).

Volatile organic contaminants can be present in the soil in four different phases, namely adsorbed to mineral soil particles and organic material, in the dissolved phase (dissolved in the soil moisture or in the groundwater), and as vapour in the unsaturated zone in the soil gas. Each phase is related to the other phases via the equilibrium equations.

The non-aqueous liquid phase is only present if the water solubility (S (mol/m^3)) of the contaminant and the vapour pressure (P (Pa)) are exceeded. As water solubility and vapour pressure are dependent on temperature, the presence of this phase is not only dependent on concentration but also on temperature.

The equilibrium relation between adsorption of contaminants to soil and the contaminant dissolved in the water can be expressed by different equations, such as the Freundlich isotherm:

$$q_s = K * C_l^n \quad n \leq 1 \tag{21.1}$$

with:

q_s= amount of adsorbed contaminant (mg/kg$_{\text{dw}}$)
C_l = concentration dissolved contaminant (mg/m^3)
K = constant
n = constant

where the value of n influences the dimension of K. The Freundlich isotherm has no adsorption maximum, which is in contrast to reality. Nevertheless, this relative simple equation is often used. If $n = 1$, a linear relation is assumed between the amount of adsorbed contaminant and the concentration of the contaminant dissolved in water:

$$q_s = K_d * C_l \tag{21.2}$$

with:

K_d= soil-water partitioning coefficient ($\mathrm{m}^3/\mathrm{kg}_{\mathrm{dw}}$)

This equation may only be applied in situations with relatively low concentrations of contaminants.

For adsorption of contaminants, the soil organic matter often plays a dominant role, therefore K_d is often substituted by K_{om} in the literature:

$$K_d = f_{om} * K_{om} \tag{21.3}$$

with:

K_{om}= organic matter-water partitioning coefficient ($\mathrm{m}^3/\mathrm{kg}_{\mathrm{om}}$),
f_{om}= fraction organic matter in soil ($\mathrm{kg}_{\mathrm{om}}/\mathrm{kg}_{\mathrm{dw}}$).

In the literature, several relations for the value of K_{om} can be found. Relations of K_{om} and the octanol-water partitioning coefficient (K_{ow}) or solubility (S) can be found as follows:

$$\log K_{om} = a * \log K_{ow} + b \tag{21.4}$$

$$\log K_{om} = c * \log S + d \tag{21.5}$$

with:

K_{ow}= octanol-water partitioning coefficient (–),
S = solubility ($\mathrm{mol/m}^3$).

In these equations different values are found for the constants a, b, c and d, for different contaminants.

The equilibrium relation between the dissolved phase and the vapour phase is described by Henry's Law:

$$P = H * C_l \tag{21.6}$$

with:

P = vapor pressure (Pa),
H = Henry coefficient (Pa m^3/mol).

This relation between the concentration in the gas phase and the concentration in the liquid phase can also be described by the dimensionless Henry coefficient (K_H), which is similar to the often used distribution coefficient m in Chemical engineering:

$$K_H = m = \frac{C_g}{C_l} \tag{21.7}$$

with:

K_H = dimensionless Henry constant (–),
m = distribution coëfficiënt (–),
C_g = concentration contamination in the gas phase (mol/m^3).

The dimensionless Henry coefficient can be derived from the ideal gas law as:

$$K_H = m = \frac{H}{R * T} \tag{21.8}$$

with:

R = gas constant (J/mol/K),
T = temperature (K).

In nearly all cases, the value of the Henry coefficient is constant, therefore the Henry coefficient can be calculated from Eq. (21.8) as the ratio of the maximal vapour pressure (P_s) and the water solubility (S):

$$H = \frac{P_s}{S} \tag{21.9}$$

with:

P_s = saturated vapour pressure (Pa),
S = water solubility (mol/m^3).

The dimensionless Henry coefficient can also be calculated as:

$$K_H = m = \frac{C_g^*}{S} \tag{21.10}$$

with:

$C^*{}_g$ = maximum concentration in gas phase (mol/m^3).

Solubility in water and the maximum concentration in the gas phase are dependent on the temperature.

The equilibrium relation between the amount of adsorbed contaminants and the contaminants in the gas phaseis comparable to Eq. (21.2):

$$q_s = K_d' * C_g \tag{21.11}$$

with:

$K'{}_d$ = soil-gas partitioning coefficient (m^3/kg dm)

Combining Eqs. (21.2), (21.7) and (21.11) leads to:

$$m = \frac{K_d}{K'_d} \tag{21.12}$$

The value of the solid-gas partitioning coefficient is dependent on temperature as well as moisture content in the unsaturated zone. When the moisture content is reduced in the unsaturated zone, the percentage of the absorbtion site susceptible to contamination at the soil surface will increase and therefore the value of K'_d will increase. For biological degradation of contaminants in the unsaturated soil, a minimal moisture content is necessary, which is comparable to about 20% of the pore volume (Koorevaar et al. 1983). At these conditions no strong increase of the K'_d value was found for several contaminants.

From the equilibrium equations above it becomes clear that, although no specific information is known about the soil characteristics, a good indication can be obtained from the physical/chemical properties of the organic contaminants on the preferential position of the contaminants in the contaminated site. As a consequence, regulators and site operators can already determine from initial site samples whether pure product can be expected.

In Fig. 21.1 an overview of the relations between the different phases is presented.

From the equations one can determine if the largest fraction of the contaminant is present in the liquid, the gas, the solid phase or as pure product. Regulators and site operators can preselect with these equations a specific remediation technology, as most technologies are directed to remediate only one specific phase in the soil system. If pure product can be expected, source treatment technologies are of interest. If high concentrations of vapour or dissolved concentrations are high, the In Situ technologies for treatment in the path of spreading are of interest. If a receptor is at risk via soil gas or groundwater contamination, specific technologies can be selected.

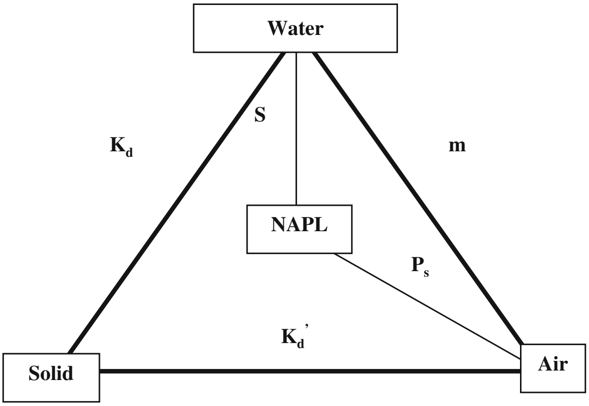


Fig. 21.1 Relations between four different phases in which organic contaminants in soil can be present

21.2.1.2 Limiting Environmental Factors

Except for insight into the distribution of contaminants over the four phases in a soil system, the kinetic parameters play the most important role in selection of a proper remediation technology. The main factors that affect the rate of the soil remediation processes are temperature, moisture, permeability of the soil, presence of an electron acceptor like oxygen, nutrients, chemical structure of the contaminant, and availability.

Temperature

A temperature increase will speed up physical/chemical processes in which soil vapour pressure and solubility of contaminants play an important role. However, in North Western Europe the average soil temperature is about 10–15°C. As temperature increase is mainly dependent on the high heat capacity of water, an antropogenically induced temperature increase of the soil is often too expensive. Due to adaptation of micro-organisms which are naturally present in soil, relatively high conversion rates for organic contaminants can be observed at temperatures from 10 to 15°C. When the temperature of groundwater is increased by more than 10°C, a shift in the microbial population will occur. After adaptation, the maximum biodegradation rate can be increased by a factor of two for each ten degrees of temperature increase, according to the Arrhenius equation. However, many microorganisms in soil and wastewater treatments systems have an optimal temperature range from the psychrophilic to the mesophilic range of about 30–35°C (Schlegel and Schmidt 1985).

Water

Biological degradation occurs only via the water phase. This means that sufficient water needs to be present. However, at a high moisture content, insufficient oxygen transport to the micro-organisms may limit the aerobic biological degradation. Suspended systems, such as bioreactors, have a very high water content, but by active aeration optimum degradation can still be ensured. In landfarming systems in sandy soil, moisture content values of 10–20% are maintained. For anaerobic biological conversions the high water content has no effect.

Permeability of Soil Water

The permeability of soil for the water (K_l) is especially of importance for the biological treatment of the saturated zone. The permeability can be calculated with Darcy's law (Koorevaar et al. 1983):

$$\frac{Q_l}{A} = K_l * \frac{\delta p}{\delta s} \tag{21.13}$$

with:

Q_l = liquid flow (m^3/s)
A = stream through surface (m^2)
K_l = permeability ($m^2/(Pa.s)$)
$\delta p/\delta s$ = pressure drop (Pa/m).

The permeability can also be expressed by an intrinsic permeability, K_{il}. This intrinsic permeability is dependent on the granule size of the soil matrix, uniformity of the soil, porosity and the moisture content. The value of K_{il} is a physical property and is independent of the water flow and pressure differences.

$$K_{il} = \eta_l * K_l \tag{21.14}$$

with:

K_{il} = intrinsic permeability (m^2)
η_l = dynamic viscosity of water (kg/(m.s)) = (Pa.s).

The dynamic viscosity of water is $1 * 10^{-3}$ (Pa.s). For high permeable soils the value of K_{il} is $1.2 * 10^{-12}$ m^2(= 1.2 Darcy) and for low permeable soils $1.2 * 10^{-14}$ m^2 (= 0,012 Darcy). The intrinsic permeability is often given in cm^2 or in Darcy (1 Darcy = $1 * 10^{-8}$ cm^2). The value of the permeability of water in soil can vary by one order of magnitude at the same location (Hinchee 1994). At very low permeability's like in clayey or loamy soils In Situ treatment by liquid transport as in pump and treat approaches is limited.

Permeability of Soil Gas

The permeability of air (K_g) is of especial importance for bioremediation in the unsaturated zone. In analogy with Darcy's Law for liquid transport, the gas transport in soil can be described by:

$$\frac{Q_g}{A} = K_g * \frac{\delta p}{\delta s} \tag{21.15}$$

with:

Q_g = gas flow (m^3/s)
A = stream through surface (m^2)
K_g = permeability ($m^2/(Pa.s)$)
$\delta p/\delta s$ = pressure drop (Pa/m).

The permeability of soil gas can also be described by an intrinsic permeability in analogy with the permeability of water. The value of K_g is a physical parameter depending on soil and is independent of the injection of airflow and pressure drop.

$$K_{ig} = \eta_g * K_g \tag{21.16}$$

with:

K_{ig} = intrinsic permeability (m^2)
η_g = dynamic viscosity gas (kg/(m.s)) = (Pa.s).

The dynamic viscosity of soil air is $20 * 10^{-6}$ (Pa.s), which is 50 times lower than for water. The soil air permeability is often expressed in cm^2 or in Darcy (1 Darcy = $1 * 10^{-8}$ cm^2). In analogy with the permeability for water the permeability of soil gas can vary one order of magnitude at the same location (Hinchee 1994). At low gas permeabilities such as found in clayey and silty soils, applicability of soil vapour extraction and bioventing is limited.

Some values of the intrinsic permeability of air, K_{ig}, for different soil types are given in Table 21.1.

Table 21.1 Intrinsic permeability of air, K_{ig}, for different soil types (Norris et al. 1994)

Soil type	K_{ig} (Darcy)
Coarse sand	100–1000
Medium coarse sand	1–100
Fine sand	0.1–1
Silt/clay	<0.1

Electron Acceptors

Especially in biodegradation, the presence and role of electron acceptor donors is evident, as is demonstrated in Fig. 21.2 (Middeldorp et al. 2002). For deep anaerobic conditions with $CO_2$and sulphate as electron acceptors, reductive dechlorination

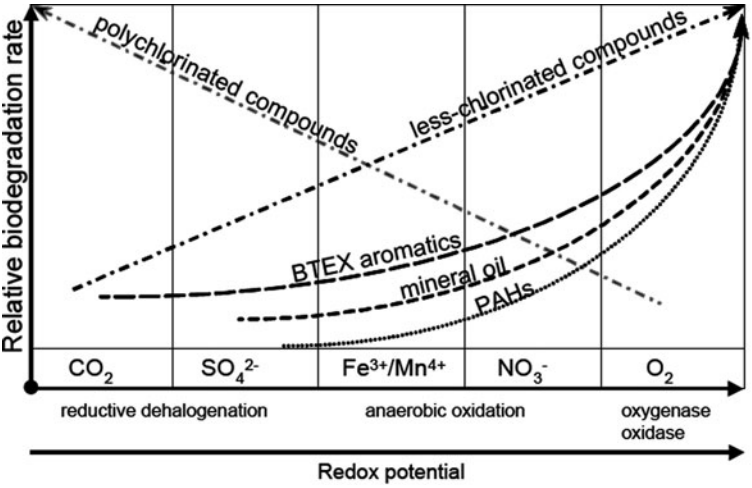


Fig. 21.2 Biodegradation rates and dominant degradation mechanisms for different classes of contaminants under various redox conditions (Middeldorp et al. 2002)

is the most rapid process to occur, whereas for the biodegradation of TPH, PAH and the aromatics benzene, toluene, ethylbenzene and xylene maximal biodegradation rates occur under aerobic conditions. For biodegradation under aerobic conditions, the supply of sufficient oxygen is important. For In Situ bioremediation in the water saturated zone, the maximal oxygen concentration is approximately 8 mg O_2/L_{water} (20°C) (= $0.25 * 10^{-3}$ (mol/L), whereas the vadose zone contains a maximum of 300 mg O_2/L_{air} ($9.375 * 10^{-3}$ (mol/L). Therefore one liter pore volume in the unsaturated zone may contain 37.5 times more oxygen than the saturated zone. In case of low permeable soils it is reported that the use of air as a carrier for oxygen is a thousand times more effective than the use of water (Wilson and Ward 1986).

Bioaugmentation

In general, biomass will be formed during the biological soil remediation process. In principle all types of micro-organisms are everywhere, as was stated by early microbiologists (Beijerinck et al. 1940). However, for relatively new and hard to degrade xenobiotics like MTBE, the inoculation with enriched microbial populations may speed up the time needed for the total bioremediation of a specific site (Salanitro et al. 2000). Many studies on bioaugmentation with specific micro-organisms have been published (Hinchee et al. 1995), but for regular observed contaminants like TPH, PAH and chlorinated solvents as such bioaugmentation is not essential.

Nutrients

To stimulate the growth of biomass, sufficient nutrients (such as nitrogen) should be supplied. Nitrogen may be supplied to the system as NH_3. Other elements which are needed for the growth of microorganisms are sulphur (S) and phosphorus (P). In the past, the growth of biomass in chemostats was studied based on mass balancing principles from chemical engineering. At that time, general element formulas were developed for biomass which include all relevant elements essential for growth of micro-organisms. A common formula for biomass is (Roels and Kossen 1978):

$$CH_{1.8}O_{0.5}N_{0.2}S_{0.0046}P_{0.0054.}$$

In the balance for the biodegradation of organic contaminants, S is often included as sulphuric acid (H_2SO_4), and P as phosphoric acid (H_3PO_4). In practice, N, S and P are added in order to create an optimum environment for biodegradation. In laboratory experiments in which the degradation of, for example, TPH is investigated in the absence of soil in order to determine the maximum conversion rate, also trace elements are added. In practice, such elements are usually present in sufficient amounts in the soil, therefore not leading to limiting conditions for biodegradation. Nutrients seem to be present in soil in sufficient quantities, as several biodegradation studies showed no improvement of biodegradation rates after addition of nutrients (Miller et al. 1991; Norris et al. 1994).

Chemical Structure

Here emphasis is on mobile compounds like chlorinated ethenes and TPH remediation. In the anaerobic conversion of PCE to ethene nearly always cis-DCE is found as intermediate (Kennedy et al. 2006). Therefore, if trans-DCE is present at a site biodegradation is inhibited (Agteren et al. 1998). For TPH remediation the effect of chemical structure on biodegradation is largely responsible for the incomplete degradation of TPH, which consists of over 500 different components. Biodegradation rates of the various compounds decrease in the following sequence: alkanes > branched alkanes > aromatics (low molecular weight) > cyclic alkanes (Geerdink et al. 1996).

For Hydrophobic Organic Compounds (HOC) the chemical structure plays an important role in their persistence to biodegradation (Agteren et al. 1998).

When chemical structure leads to persistence, regulators and site operators will most often use physical techniques to control a specific site.

Toxicity

Several organic contaminants may be toxic for micro-organisms. The effects of many organic contaminants seem to be related to the octanol-water partitioning coefficient (K_{ow}). The value of this coefficient gives the ratio of a contaminant between the organic fraction (octanol) and water. Many organic contaminants have very high values of this partitioning coefficient, therefore often the log K_{ow} is given for organic contaminants. In general, it is found that organic solvents with log $K_{ow} > 4.0$ do not inhibit microbial degradation. Organic solvents with log $K_{ow} < 2.0$ show a negative impact on micro-organisms (Laane et al. 1987). In specific cases contaminants with $K_{ow} < 2.0$ can be degraded, whereas contaminants with a high log K_{ow} sometimes may lead to effects. In Table 21.2 some log K_{ow} values are given for non chlorinated organic contaminants.

For chlorinated ethenes the K_{ow} decreases during conversion from PCE to TCE, cis-DCE and VC, therefore the toxicity in fact increases. However this is only a temporary effect as the final product is the harmless ethane (Agteren et al. 1998).

Table 21.2 Log K_{ow} values are given for non chlorinated organic contaminants

Contaminant	Log K_{ow}
Dioxane	−1.1
Acetone	0.23
Pyridine	0.71
Benzene	2.0
Toluene	2.4
Xylene	3.1
Diethyl phthalate	3.3
diphenyl ether	4.3
Decane	5.6
Tetradecane	7.6
Dioctyl phthalaat	8.8

Source: Laane et al. (1987)

Except this general pragmatic approach as first estimate of toxicity to microorganisms, the effect of toxicity can be rather complex. For example the cometabolic aerobic conversion of 1,1 cis-DCE will lead to formation of toxic intermediates and therefore bioremediation will cease (Oldenhuis et al. 1991).

Availability

In addition to the above factors, the bioavailability of contaminants plays an important role, especially in attaining low residual concentrations in treated soils.

For soil remediation most of the factors above can be manipulated to maximize conversion or removal rates. However, as limited availability is believed to be the result of a long aging procedure in which especially the soil organic matter, but also the surface area of especially small sized particles like clay particles play a role, the limited availability seems to be a principle obstacle for In Situ remediation to negligible concentrations.

From a risk perspective only the available fraction of contaminants impose a risk to the soil ecosystem (see Chapter 16 by Hodson et al., this book). However, from a remediation perspective, only the available fraction can be removed or remediated. The combination of these perspectives is leading to new approaches in legislation in which residual concentrations after soil remediation are accepted as end points, as long as the availability of the contaminants with time is negligible. On the other hand, most legislations are still based on concentrations rather than on mass fluxes that are transported from the soil to the surrounding water in which they pose a risk for uptake in the food chain. At present several techniques are described to determine the available fraction of Hydrophobic Organic Compounds, which helps to obtain insight in the potential risk of a contaminant (Cornelissen et al. 1998; Cuypers et al. 2001; Reid et al. 2000). However only recently insight was obtained under what conditions contaminants will be released from soils and sediments (Smit 2009; Smit et al. 2008). These release studies showed that turbulent conditions like flooding and transfer of sediment into depots will temporarily lead to high contaminant fluxes to the environment.

21.2.2 In Situ Technologies

For In Situ technologies one can discriminate between source, path and receptor oriented technologies (Fig. 21.3). These different measures are often linked to different parts of legal instruments such as in Europe the Groundwater Directive, the Water Framework Directive, and the Soil Thematic Strategy (Chapman et al. 2008).

21.2.2.1 Source Oriented In Situ Technologies

Physical and chemical technologies are especially well suited to treat sources with pure products of contamination, as bioremediation can only function in an aqueous environment and contaminant concentrations should be below toxic concentrations.

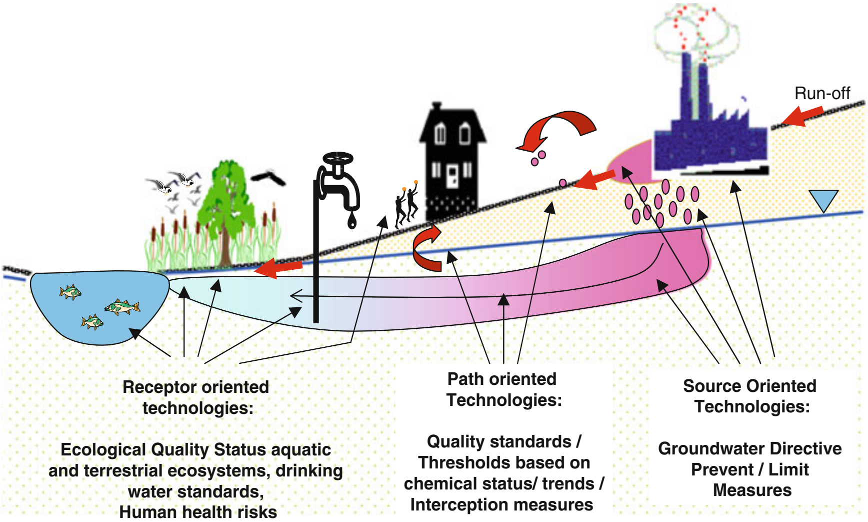


Fig. 21.3 Source-path-receptor approach and source, path and receptor oriented technologies

Physical techniques like removal of a floating layer by pumping are described in the literature, e.g., *pump and treat*, *slurping or bioslurping* (Alleman and Leeson 1997). Especially, *removal of floating layers with pure product* is very efficient as further spreading in the environment will be avoided. However, pump and treat technologies are less efficient when only the dissolved fraction in the groundwater can be removed (Hoffman 1998). In this situation the use of steam injection (Smart 2005), surfactant enhanced cleaning (Hayworth 1997), or co-solvent flushing (Ramakrishnan et al. 2005) may be useful. After the first applications of pump and treat in the early 1990s, it became clear that continuous pumping was often not cost-effective, as a large fraction of contaminants was tightly bound to the soil matrix, especially for contaminants with a high K_d value as is shown in Fig. 21.1. Thereafter intermittent pumping of contaminated groundwater was applied as an alternative.

Physical techniques for volatile contaminants were developed in a similar way, as continuous stripping of soil gas was soon followed by intermittent soil gas removal, like venting and intermittent venting. One of the disadvantages of severe stripping was that, after treatment, catalytic thermal treatment or biofiltration had to be applied to the stripped soil gas. Shortly thereafter, it became clear that biodegradation could also be applied by the endogenous microbial population in the vadose zone, and technologies such as *bioventing* were introduced (Malina et al. 2002). The extra gas that was introduced in the soil was not specifically directed to strip the volatile contaminants, but the introduced oxygen that came with the air was used for enhanced biodegradation for BTEX contaminant as well as for other volatile TPH compounds. For the saturated zone, volatile contaminants can be removed by air sparging, however in the early 1990s, risks of spreading of contaminants by this technique were regarded as a disadvantage. (Johnson et al. 1993).

In conclusion the source oriented physical techniques for volatile contaminants were shown to be less effective and were replaced by more effective treatments in the path of these contaminants.

Chemical treatment by *In Situ Chemical Oxidation* (ISCO) has been described at field scale since about 2002 (Hartog et al. 2004; Plaisier et al. 2003). In Situ Chemical Oxidation (ISCO) involves the introduction of chemical oxidants into the soil to destroy organic contaminants (e.g., chlorinated organic solvents or fuels). A wide variety of chemical oxidant types exist. Most frequently used for ISCO are hydrogen peroxide (H_2O_2, often in Fenton's Reagent mixture with ferrous iron), permanganate (as $KMnO_4$ or $NaMnO_4$) and ozone (O_3). Recently, persulfate ($Na_2S_2O_8$) has emerged as an alternative oxidant for ISCO. These chemical oxidants are stronger than those that occur in biological oxidation reactions. Due to their higher oxidative strength they can oxidize a wider variety of contaminants at a faster rate. In addition, complete oxidation of the target contaminant generally occurs without the formation of potentially harmful intermediate contaminants. Due to the non-biological nature of the oxidation process, high contaminant concentration, including NAPL phases, can be oxidized unhampered by toxicity effects. Despite the wide range of reaction characteristics for the multitudes of chemical oxidant types available, general key issues relating to the sustainability and cost-effectiveness of ISCO can be identified. First the loss of oxidants through reactions with the natural soil oxidant demand (e.g., components such as organic matter and iron sulphides) should be minimized. A successful ISCO application therefore requires optimal oxidant loading (dose concentration and delivery) for a particular contaminated soil system to maximize cost-effectiveness and minimize soil disturbance (Haselow et al. 2003; Mumford et al. 2004; Nelson et al. 2001). The application of strong oxidants may lead to the unwanted mobilization of metals, as due to the oxidation of organic matter the binding capacity is expected to decrease and also the formation of contaminants as byproducts might occur (Crimi and Siegrist 2003, 2004). Other researchers point to potential reduction of soil permeability as the faster reaction rates promote the accumulation of reaction precipitates and gas formation (Lee et al. 2003). Also some research groups are concerned about the recovery of biological soil functions (Ecosystem Services), including Natural Attenuation capacity, after ISCO application (Christ et al. 2005; Sahl et al. 2007). However, others promote the combination of ISCO and biodegradation. The rationale for this is to reduce the amount of chemical oxidants and to aim at partial oxidation followed by biodegradation of the intermediate contaminants. Further monitoring procedures need to be developed to improve the monitoring of the remediation process (Cave et al. 2007).

Most In Situ remediation technologies focus on organic contaminants, as in general for heavy metals there are no other options than mobilization or immobilization of metals. For In Situ mobilization of metals, only a limited number of studies have been performed. In a sandy soil, Jansen et al. (2004) were able to mobilize zinc in a sandy soil at pH= 4. The acidic effluent downstream of the site was treated by microbial sulphate reduction, leading to precipitation of ZnS in a relatively small volume to concentrate the waste stream.

A similar immobilization was performed with the Zn^{2+} contaminated groundwater at a zinc smelter in Budel, the Netherlands. Here, the controlled process of sulphate reduction was used to concentrate the heavy metal as ZnS in a small volume that could be taken up in the zinc melting process thereafter (Janssen and Temminghoff 2004).

21.2.2.2 Path Oriented In Situ Technologies

Between the source of contamination and the receptor, the plume or path of the contaminants is found. In this area, which can have dimensions varying from meters to kilometres, several treatment methods have been developed.

"Early path" In Situ technologies include *the funnel and gate system* (Starr and Cherry 1994). Here, emphasis was on the physical phenomena related to the groundwater flow, making use of the natural groundwater flow instead of the pump & treat methods that were popular at the start of In Situ treatment. The installation of cut-off walls in combination with In Situ bioreactors in the gate of the system led to a less energy consuming treatment.

In later years the concept was extended to all types of biological active screens or *bioscreens* (Fig. 21.4) in which attention was mainly focused on the biological

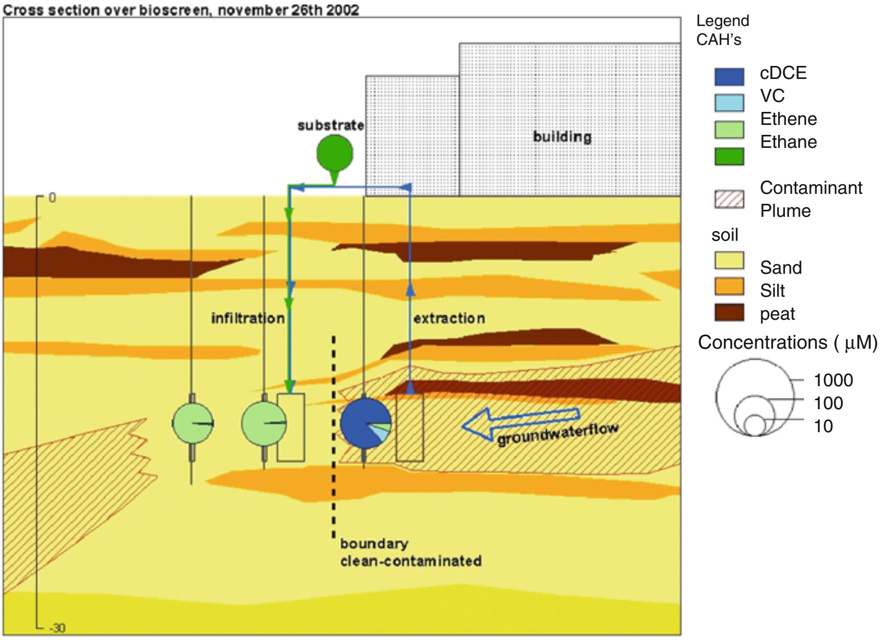


Fig. 21.4 An effective In Situ bioscreen at a site contaminated with chlorinated solvents in the Netherlands

processes (Hoekstra et al. 2005; Middeldorp et al. 2002). Also the term *Permeable Reactive Barriers* (PRBs) became known in which also chemical conversions were included, especially for chlorinated solvents by Zero Valent Iron (ZVI) (Gu et al. 1999; Scherer et al. 2000). A wide variety of PRBs is applied at field scale (Van Nooten et al. 2008) and in recent years the physical design such as in the Biological Fence (Langenhoff et al. 2002) is changed into more extensive methods like Active Plume Management, Enhanced Bioremediation (Van Heiningen et al. 1999), and Natural Attenuation (Illman and Alvarez 2009). All these methods aim at stimulation and monitoring the biodegradation of contaminants in soil and groundwater. See Chapter 22 by Peter et al., this book, for a detailed analysis of Natural Attenuation. Recently, new approaches have been applied in practice using Natural Attenuation (NA) at reactive interfaces, such as between groundwater, sediment and surface water (Fig. 21.5) and between groundwater and air (the vadose zone of soils). Generally speaking, the subsoil and groundwater have a lower redox potential and oxygen level than aerobic surface waters and the water-unsaturated upper layer of soils. Here, contaminants that migrate by diffusion or convection from the anaerobic compartment (groundwater, sediment) to the aerobic receptor (vadose zone/indoor air, surface water), may become biodegradable and can be removed from air and water by natural biodegradation processes (De Weert et al. 2008; Middeldorp et al. 2005), (http://www.vito.be/sedbarcah), (Abreu and Johnson 2006; Abreu et al. 2009; Devaull 2007). More laboratory, field and modelling research is required for the safe and effective use of these NA-interface approaches.

Regularly updated information on path oriented In Situ technologies is presented at several websites such as Eurodemo (http://www.eurodemo.info/).

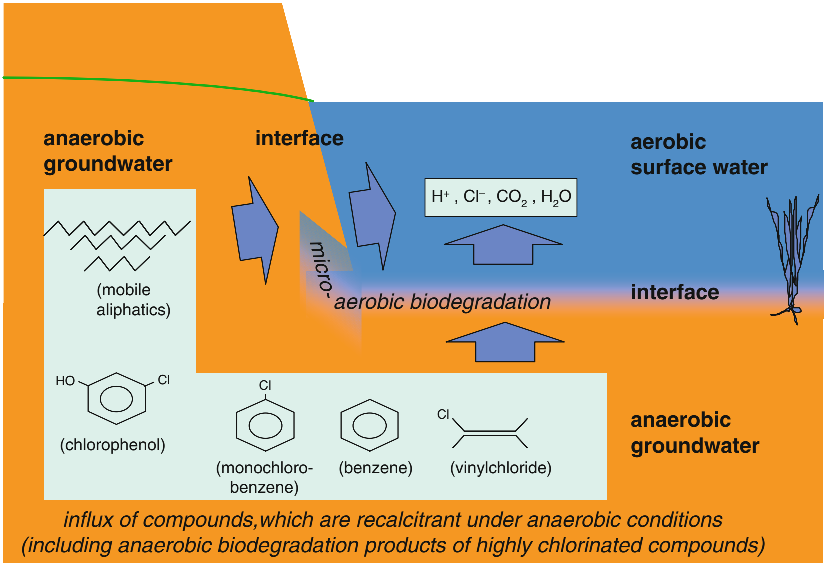


Fig. 21.5 Natural Attenuation at the groundwater-surface water interface

21.2.2.3 Receptor Oriented In Situ Technologies

The "receptor In Situ Technologies" aim at the direct protection of the receptor, like a well for human drinking water supply or a groundwater well for cattle, or a surface water body. Combinations of biological and, physical techniques are often applied at this receptor level, and often concern highly mobile contaminants that are not, or only slowly, biodegraded in natural systems.

Drinking Water

A typical example is methyl tertiary butyl ether (MTBE), an oxygenation additive in gasoline. Although the International Agency for Research on Cancer does not classify this compound as a human carcinogen, the required concentration in water for MTBE is below 5–15 μg/L if it is used for drinking water preparation, as it can be tasted at these low levels in water (Fischer et al. 2005) Although some biodegradation is reported (Haggblom et al. 2007), many groundwater extractions for drinking water are affected by MTBE. Generally, the extracted water is treated using various treatment methods. Shih et al. (2003) studied sorption by different granular activated carbons (GAC) and demonstrated that competitive adsorption of co-contaminants (like BTX) and natural organic matter in surface or groundwater can greatly reduce MTBE removal efficiencies. Adsorption with zeolites and treatment with advanced oxidation techniques using ozone and peroxide is another promising technology, currently in development (http://www.sense.nl/research/1869).

Chemical oxidation using Fentons Reagent was studied in batch reactors with anoxic groundwater by Burbano et al. (2005) who showed highest MTBE removal at acidic pH, which strongly reduced efficiencies at neutral conditions. Although MTBE concentrations decreased with 90–99%, mineralisation occurred only for 30–40% leaving tert-butyl formate (TBF), tert-butyl alcohol (TBA), acetone and methyl acetate as the major intermediate contaminants. Thus a second chemical or biological "polishing step" would be needed to complete the mineralisation.

Surface Water

Another set of technologies are oriented at contaminated sediments with the surface water system as the receptor. As sediments release their contaminants to surface water, dependent on river flow characteristics (Smit 2009), they will affect the quality of the surface water as regulated for example in the EU in the Water Framework Directive (http://ec.europa.eu/environment/water/water-framework/index_en.html).

Koelmans et al. (2006) showed that soot and other black carbon particles present in sediments strongly bind and accumulate diffuse hydrophobic contaminants. These and other researchers are currently transferring this phenomenon into treatment technologies. Hereto, activated carbon and other sorptive materials are added to the upper layer of the contaminated sediment to reduce emissions of strongly sorbing contaminants, such as Polycyclic Aromatic hydrocarbons (PAH's) and Polychlorinated Biphenyls (PCBs). Cho et al. (2009) demonstrated short and long term reductions of up to 90% in bioaccumulation of PCBs in marine clams by amendments of 2% activated carbon to contaminated sediments. A similar test and observation was done with PAH contaminated urban soils and sediments (Brandli et al. 2008). These technologies are currently being tested for various field situations and can be expected to become applied at large scale in the near future.

In the Netherlands, an expert team evaluated possible risks of immobile contaminants including metals in sediments and submerged soils. This expert team concluded that for each case an insight into the processes involved needs to be established before adequate Risk Assessments and emission reduction measures (such as reactive amendments and capping) can be applied (Schipper et al. 2009). In general, capping should be used in environments where the long-term physical integrity of the cap can be maintained, and environments with low turbulence are generally desired for In Situ capping projects. The potential severity of the environmental impacts associated with cap erosion and potential dispersion of the sediment contaminants in an extreme event should determine the level of protection against erosion (http://www.epa.gov/glnpo/sediment/iscmain/).

21.3 Integration of In Situ Technologies in Risk Management

21.3.1 Risk Management Concepts and Frameworks

A distinction can be made between Risk Assessment and Risk Management (Brils et al. 2008; see Chapter 1 by Swartjes, this book; Van Leeuwen and Vermeire 2007). Risk Assessment is a science based method using site data and models identifying and calculating hazard, exposure, and effects at the receptor. By applying uncertainty and probability models the risk, i.e., the probability for damage to a receptor in time and space, is determined. Risk Management concerns possible responses to risks by actions. These can be interventions such as remediation and interception measures at the source, path or receptor, or can be in the form of alterations in the functional use of a contaminated area. In addition, Risk Management decisions are often based on considerations that are influenced by other factors (e.g., economic, political, public perception). Van Leeuwen and Vermeire (2007) defines therefore a third component, namely *Risk Communication*, taking the societal context, involvement of stakeholders, and the current policy on Risk Assessment and evaluation into account.

Vegter et al. (2003) and others introduced the concept of *Risk Based Land Management* (RBLM), integrating two key decisions for the remediation of contaminated land:

1. the time frame: this requires an assessment of risks, priorities, and the longer-term effects of measures;
2. the sustainability of the intervention measure or change in functional use: this requires an assessment of overall benefits, costs and environmental side effects, the value and condition of the land, public and local community perceptions and other issues.

In order to arrive at sustainable solutions these two key decisions need to balance three basic principles: (i) fitness for use, (ii) protection of the environment and iii) long-term care (Vegter 2004). Although RBLM is now acknowledged by many authorities (www.commonforum.eu) as the approach to follow, many site remediation programmes still find major obstacles in applying this in practice, for example lack of urgency, lack of money, difficulties in recycling slightly contaminated soil (Sorvari et al. 2009). Moreover, the outcome of quantitative assessments of costs and benefits, is strongly related to politically, socio-economically and ethically defined pre-assumptions to be chosen as starting points for the cost-benefit analyses itself (Van Wezel et al. 2007). Major developments are therefore still needed to further improve and bench mark methodologies for sustainable risk based decision making in soil, groundwater and sediment management and remediation.

A useful concept for this can be the DPSIR model (Brils et al. 2008). This concept treats the environmental management process as a feedback loop controlling a cycle of Driving forces (D), Pressures (P), States (S), Impacts (I) and Responses (R). Economic urban, agricultural, and industrial activities (Driving forces) lead to increasing Pressures on the natural environment by the use of natural resources and/or emissions to (ground) water, soil, and sediment. That changes the State of these environments in quantity and quality. Response measures can be implemented at any of the D, P, S, or I levels, to prevent, limit or mitigate the impacts to acceptable levels. Thus, the DPSIR framework links scientifically based Risk Assessment to decision making, management and policy. Building on DPISR the European project RISKBASE (Brils et al. 2008) presents a modified framework for Risk Management (Fig. 21.6).

Socio-economic and global change, including climate and land use change, are taken as an autonomic driver-pressure sequence influencing the biophysical system (like a contaminated site) to which the social system responds. At one side, people observe and gain understanding of the biophysical system. At the other side, people respond by policy, management, and public actions in terms of measures leading to the sustainable maintenance of the biophysical system. It is an adequate

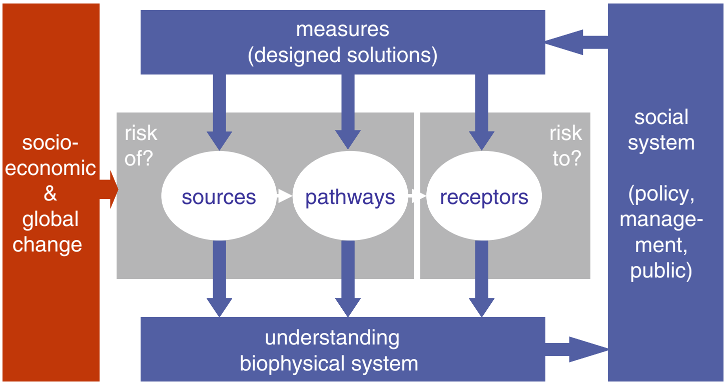


Fig. 21.6 RISKBASE Risk Management framework (Brils et al. 2008)

structure to use for soil and sediment remediation, since this framework is just like the technological measures (Section 21.2) oriented at sources, pathways and receptors. Moreover, it also offers possibilities to further develop the sustainability part of soil and sediment management. The biophysical soil-sediment-water system can be defined as one of the receptors providing goods and services to society. In this way the framework connects to the concept of Ecosystem Services which is currently considered as a new leading principle in sustainable environmental management (see Chapter 13 by Swartjes et al., this book), addressing not only soil and sediment remediation but also river basin management, climate adaptation and mitigation, and the sustainable use of the subsurface.

21.3.2 Risk Management Application

To make risk information easily accessible for decision making, one often chooses a generic tiered approach using traffic light models (IRGC 2007), distinguishing a

- green mark for acceptable risks: no further actions are needed;
- yellow mark for acceptable: actions are needed to reduce risks over time and space, in accordance to the ALARP – "as low as reasonably possible" – principle and the BAT "best available technique" – approaches;
- red mark for acceptable risks: direct interventions have to be made by remediation actions and/or by changing the functional use of the site.

This concept has been applied for Natural Attenuation in groundwater (Sinke et al. 1998), human health risks associated with contaminated soil and groundwater (Bien et al. 2005), and emission risks from sediments contaminated with metals, PAH and hydrophobic chlorinated contaminants (such as hexachlorocyclohexanes, polycholorobenzenes, Dieldrin, and PCB's) (Malina et al. 2002).

Risk Assessment requires a minimum of data and information related to the site before it can be used in a preliminary set up for Risk Management. In practice, and especially in large areas with soil and sediment contamination, this is often complicated. Acquiring site information and building up of an understanding of all contaminant migration pathways, receptor exposure routes and uncertainties often requires time and investments, while management decisions need to be taken in much shorter time frames. A solution for these situations is to form a group of experts and stakeholders that estimates the risks and evaluates various intervention and remediation options (Malina et al. 2002; Ter Meer et al. 2008). Here, natural science based risk data, expert judgements and socio-economic arguments of stakeholders modulate the outcome of Risk Management decision. For large scale problems such as sediments in rivers and lakes, and contaminated land and groundwater at Megasites, a cyclic evolving approach is needed. In this way Risk Assessments and Management strategies can evolve with the build up of information and understanding of processes occurring at the site or region. This generally requires a time frame of years to a decade (Ter Meer et al. 2008).

21.3.3 Risk Management at Contaminated Megasites

In the EU project WELCOME (http//www.EUwelcome.nl/) the management of contaminated soil, groundwater and surface water on large (former) industrial sites has been addressed (Malina et al. 2002; Ter Meer et al. 2008). The extent of contamination at these so-called *Megasites* is so large and complex that an integrated cost-effective approach is required to reduce and mitigate the risks. An *Integrated Management Strategy* (IMS) was developed in line with the EU water framework and groundwater directives (http//www.EUwelcome.nl/kims). Especially, the prevent-and-limit paragraphs of the EU groundwater directive are addressed (Ter Meer et al. 2008). The IMS has been demonstrated by applications at four Megasite cases, namely Bitterfeld, Germany; Rotterdam Harbour Region, the Netherlands; Antwerp Harbour Region, Belgium; and Tarnowskie Gory, Poland. The IMS is aimed at guiding the user through the whole process of establishing a Megasite management plan; from the initial screening to the final definition of the remediation scenarios and long-term site management plan. It includes:

- basic strategies of the IMS;
- a manual for technical experts and environmental managers;
- examples of the implementation of the IMS at the Welcome Megasite cases;
- an overview of the technical support tools.

The IMS-manual helps the user to distinguish the areas of the Megasite with the highest risk and set up priorities for the level of risk reduction, degree of remediation and consequently, related investments on the basis of Risk Assessment for the area. This process can often minimize the costs of the management and remediation measures and maximize its effectiveness by channelling investments into the areas of the site with the highest risk. Four basic phases are distinguished in the IMS, i.e., Starting the IMS, Management scenarios, Risk Assessment, Implementation.

21.3.3.1 Starting the IMS

The overall objective of this section is the provision of all criteria needed to define a site as a Megasite, to derive the specific management tasks and objectives, and to establish a group of stakeholders and experts. An important activity is to establish a site conceptual model (see Fig. 21.7 for the Rotterdam harbour as an example) with possible migration pathways defining planes of compliance to be used in further Risk Assessment and Management.

21.3.3.2 Risk Assessment

The guideline consists of the following steps: Carry out a Megasite characterization; Define potential risk clusters; Carry out fate and transport modelling; Determine risks and derive local standards; Finalize risk clusters. In the Rotterdam case risks were especially relevant for the groundwater system directly below the harbour region, and not for the surrounding surface waters and polders. In Fig. 21.8 the area

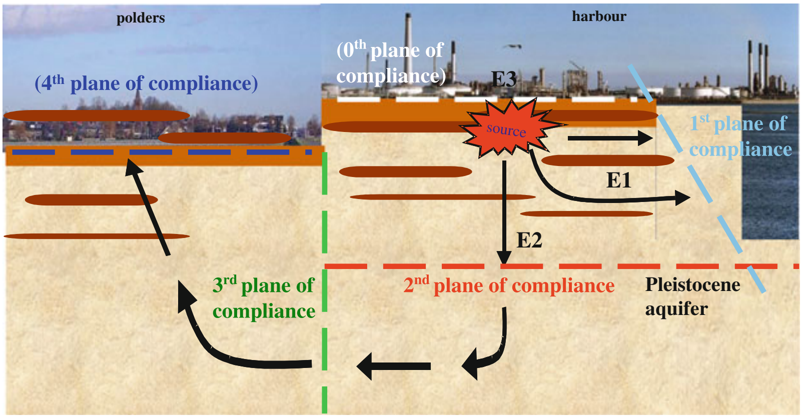


Fig. 21.7 Example of a Megasite conceptual site model; The Rotterdam Harbour example is presented here. The planes of compliance refer to conceptual boundaries that should not be affected by contamination

of aquifer of The Rotterdam Harbour (in % of total area) that exceeds the Dutch Intervention Value as a function of time is given, as an example of the results from a Megasite Risk Assessment. For the Antwerp Harbour this situation was completely different. Here, all risks were associated with emissions from shallow groundwater to surface water and not to deeper groundwater systems. This was due to the presence of the aquitard “Boomse clay” at a depth of 10 m below surface level. This shows that a thorough system understanding is required for adequate Risk Assessment and following Risk Management steps.

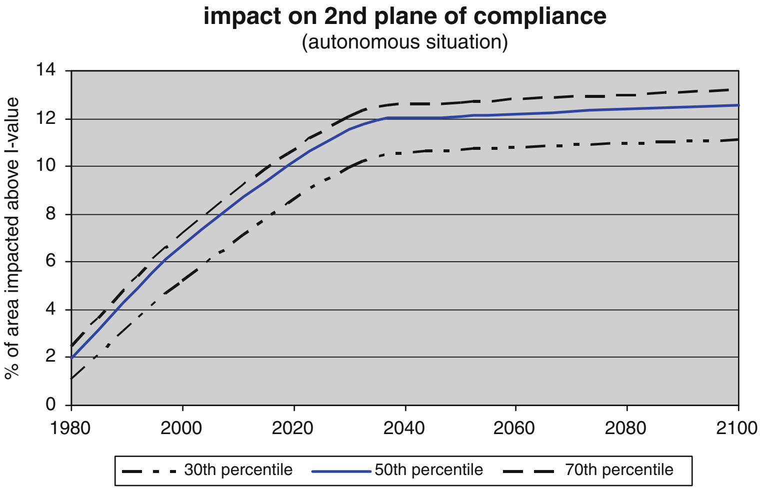


Fig. 21.8 Area of aquifer of The Rotterdam Harbour (in % of total area) that exceeds the Intervention Value as a function of time, as an example of the results from a Megasite Risk Assessment

21.3.3.3 Management Scenarios

This section helps to define management scenarios for the Megasite that are cost-effective and sustainable. This implies that in the long-term a maximum risk reduction will be achieved within the budget available. The following activities need to be carried out:

- define the feasibility of management scenarios for each cluster (including risk reduction objectives and priority);
- perform cost-efficiency and risk reduction analysis;
- prioritize and optimise management scenarios for the Megasite.

In the Example for The Rotterdam Harbour shown in Fig. 21.9 as a conceptual site model, Strategies with combinations of measures using pump and treat, In Situ remediation and Natural Attenuation were selected for further evaluation. Cost-benefit analyses also need to be performed at this stage. In Fig. 21.10 the costs and forecasted effects on groundwater quality of different scenario's are shown for The Rotterdam Harbour case.

21.3.3.4 Implementation

Once the optimal final scenario has been selected by the stakeholders, it needs to be implemented, monitored and reviewed. An implementation plan will need to be developed within the IMS to support successful site management. Part of that is a monitoring programme at the planes of compliance to control the performance of

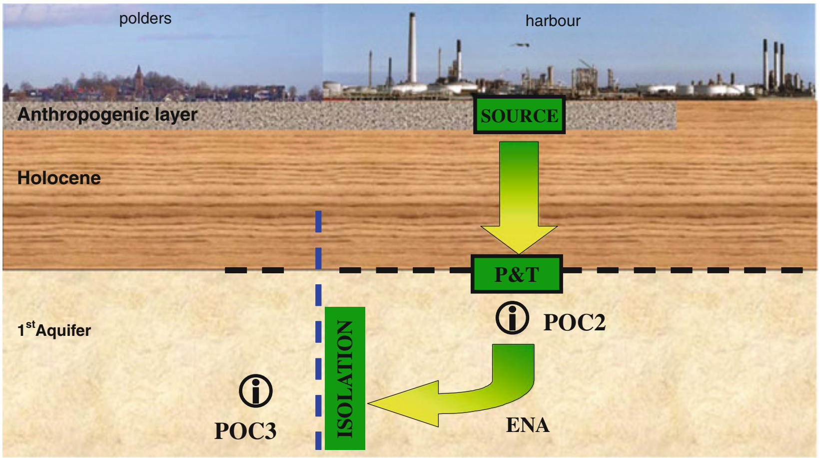


Fig. 21.9 Conceptual site model of The Rotterdam Harbour, showing a mixed strategy of measures at planes of compliances 2 and 3: P&T = pump and treat, and ENA (Enhanced Natural Attenuation; In Situ bioremediation including Natural Attenuation), as an example of a Megasite management scenario

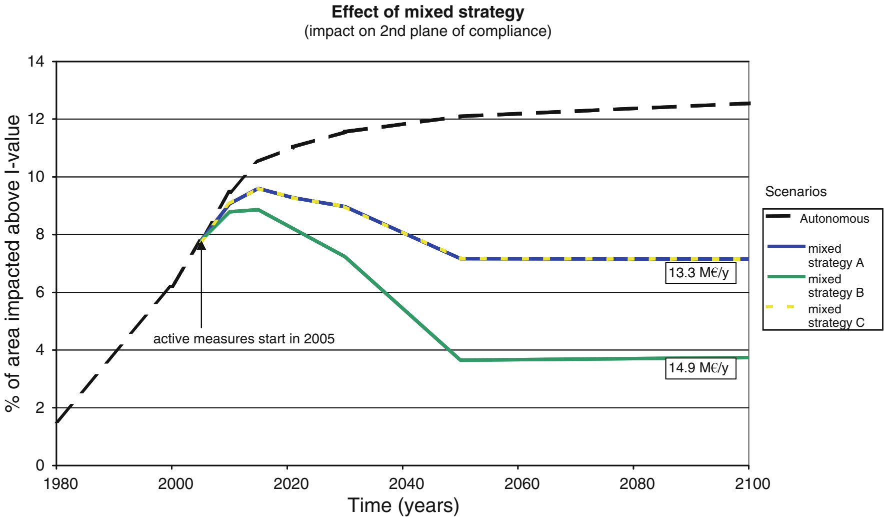


Fig. 21.10 Area of aquifer of The Rotterdam Harbour (in % of total area) that exceeds the Intervention Value as a function of time for three remediation and management scenarios, as an example of cost-efficiency assessment

the management scenario, to fill previously defined data gaps and to carry out further site investigations to reduce uncertainties in the information on which the scenario is based. Adjustments of the management plan, on the basis of new information, are foreseen in the implementation phase.

21.4 Outlook

Remediation needs to be further placed in the framework of sustainable Risk Management. More developed methods for cost-benefit assessments are needed to evaluate different remediation approaches and technologies. Conventional remediation techniques need more energy than In Situ technologies and the associated production of CO_2 has raised doubts about their overall sustainability. The introduction of green remediation approaches can significantly boost the sustainability of regional management of soil, sediment, water and groundwater quality (see Chapter 20 by Bardos et al., this book). For Megasites like the Port of Rotterdam, and urban regions, the key to this ambition lies in combining parallel goals for energy and water. For instance it is possible to effectively combine industrial water use and industrial heat surplus with the remediation of contaminated groundwater. In urban regions aquifer thermal energy storage can be combined with remediation. This will lead to both an improvement of the groundwater quality and a reduction of energy use and CO_2 emissions.

Soil, water, sediment and groundwater quality management, including Risk Based In Situ remediation, is likely to become much more efficient when placed all in an integrated holistic planning associated with river basin management (EC COM 231 2006), (EC COM 232 2006) (Gerzabek et al. 2007). It is important to note that most threats to soil identified in the Thematic Strategy on Soil Protection (EC COM 179 2002) have strong relationships with the way water is managed. Therefore in order to manage surface waters, groundwaters and sediments towards AN ecologically satisfactory state, it is essential to include the management and protection of soils in this and vice versa (Brils et al. 2008).

References

Abreu LDV, Ettinger R, McAlary T (2009) Simulated soil vapor intrusion attenuation factors including biodegradation for petroleum hydrocarbons. Ground Water Monitor Remed 29:105–117

Abreu LDV, Johnson PC (2006) Simulating the effect of aerobic biodegradation on soil vapor intrusion into buildings: influence of degradation rate, source concentration, and depth. Environ Sci Technol 40:2304–2315

Agteren MHV, Keuning S et al (1998) Handbook on biodegradation and biological treatment of hazardous organic compounds. Kluwer. ISBN 0-7923-4989-X

Alleman BC, Leeson A (1997) In situ and on site bioremediation volume 1, fourth international in situ and on site symposium. Batelle, Columbus, OH

Beijerinck MW, Van Iterson G, den Dooren de Jong LE, Kluyver AJ (1940) (In Dutch) Verzamelde geschriften Van M.W. Beijerinck: ter gelegenheid van zijn 70sten verjaardag, pp 1921–1922

Bien J, ter Meer J, Rulkens W, Rijnaarts HHM (2005) A GIS-based approach for the long-term prediction of human health risks at contaminated sites. Environ Model Assess 9:221–226

Brandli RC, Hartnik T et al (2008) Sorption of native polyaromatic hydrocarbons (PAH) to black carbon and amended activated carbon in soil. Chemosphere 73(11):1805–1810

Brils J, Barceló D, Blum WEH, Brack W, Harris B, Müller D, Négrel Ph, Ragnarsdottir V, Salomons W, Track Th, Vegter J (2008) River basin risk assessment linked to monitoring and management. Chapter 7.1 In: Quevauviller P, Borchers U, Thompson KC, Simonart T (eds) The water framework directive: ecological and chemical status monitoring. Wiley, Chichester, pp 273–287

Burbano AA, Dionysiou DD, Suidan MT, Richardson TL (2005) Oxidation kinetics and effect of pH on the degradation of MTBE with Fenton reagent. Water Res 39(1):107–118

Cave L, Hartog N et al (2007). Electrical monitoring of in situ chemical oxidation by permanganate. Ground Water Monitor Remed 27(2):77–84

Chapman A, Brils J, Ansink E, Herivaux C, Strosser P (2008) Conceptual models in river basin management. In: Quevauviller P (ed) Groundwater science and policy. An international overview, RSC, ISBN 978-085404-294-4

Cho YM, Ghosh U et al (2009) Field application of activated carbon amendment for in-situ stabilization of polychlorinated biphenyls in marine sediment. Environ Sci Technol 43(10):3815–3823

Christ JA, Ramsburg CA et al (2005) Coupling aggressive mass removal with microbial reductive dechlorination for remediation of DNAPL source zones: a review and assessment. Environ Health Perspect 113(4):465–477

Crimi ML, Siegrist RL (2003) Geochemical effects on metals following permanganate oxidation of DNAPLs. Ground Water 41(4):458–469

Crimi ML, Siegrist RL (2004) Association of cadmium with MnO_2 particles generated during permanganate oxidation. Water Res 38(4):887–894

Cornelissen G, Rigterink H et al (1998) Rapidly desorbing fractions of PAHs in contaminated sediments as a predictor of the extent of bioremediation. Environ Sci Technol 32:966–970

Cuypers C, Clemens R et al (2001) Prediction of petroleum hydrocarbon bioavailability in contaminated soils and sediments. Soil Sediment Contamin 10(5):459–482

Cuypers C, Grotenhuis T et al (2000) Rapid persulfate oxidation predicts PAH bioavailability in soils and sediments. Environ Sci Technol 34(10):2057–2063

Devaull GE (2007) Indoor vapor intrusion with oxygen-limited biodegradation for a subsurface gasoline source. Environ Sci Technol 41(9):3241–3248

Dragun J (1991) Geochemistry and soil chemistry reactions occurring during in situ vitrification. J Hazard Mater 26(3):343–364

Fischer A, Oehm C et al (2005) Biotic and abiotic transformations of methyl tertiary butyl ether (MTBE). Environ Sci Pollut Res 12(6):381–386

Gates DD, Siegrist RL (1995) In-situ chemical oxidation of trichloroethylene using hydrogen-peroxide. J Environ Eng-Asce 121(9):639–644

Geerdink MJ, Kleijntjens RH et al (1996) Microbial decontamination of polluted soil in a slurry process. J Environ Eng-Asce 122(11):975–982

Gu B, Phelps TJ et al (1999) Biogeochemical dynamics in zero-valent iron columns: implications for permeable reactive barriers. Environ Sci Technol 33(13):2170–2177

Haggblom MM, Youngster LKG et al (2007). Anaerobic biodegradation of methyl tert-butyl ether (MTBE) and related fuel oxygenates. Adv Appl Microbiol 62:1–20

Hartog N, Van Bergen PF et al (2004) Reactivity of organic matter in aquifer sediments: geological and geochemical controls. Geochim Et Cosmochim Acta 68(6):1281–1292

Hinchee RE, Fredrickson J, Alleman BC (1995) Bioaugmentation for site remediation. 3rd international in situ and on site bioreclamation symposium, vol 3. Batelle Press, Columbus, OH. ISBN AQ2 1-57477-004-7

Hoekstra N, Langenhoff A et al (2005) Enhanced microbial degradation of chloroethenes and chloroethanes in a bioscreen. In: Boshoff G, Bone BD (eds) Permeable reactive barriers, vol 298. IAHS, Oxfordshire, pp 75–81

Hoeppel RE, Hinchee RE et al (1991) Bioventing soils contaminated with petroleum-hydrocarbons. J Indust Microbiol 8(3):141–146

Hoffman AH (1998) Pump and treat rescue. Civil Eng 68:56–57

Hutzler NJ, Murphy BE et al (1991) State of technology review – soil vapor extraction systems. J Hazard Mater 26(2):225–229

Illman WA, Alvarez PJ (2009) Performance assessment of bioremediation and natural attenuation. Crit Rev Environ Sci Technol 39(4):209–270

Janssen G, Temminghoff, EJM (2004) In situ metal precipitation in a zinc-contaminated, aerobic sandy aquifer by means of biological sulfate reduction. Environ Sci Technol 38(14):4002–4011

Johnson RL, Johnson PC et al (1993) An overview of in-situ air sparging. Ground Water Monitor Remed 13(4):127–135

Kennedy LG, Everett JW et al (2006) Field-scale demonstration of induced biogeochemical reductive dechlorination at Dover air force base, Dover, Delaware. J Contamin Hydrol 88(1–2):119–136

Koelmans AAM, Jonker TO et al (2006) Black carbon: the reverse of its dark side. Chemosphere 63(3):365–377

Koorevaar P, Dirksen C et al (1983) Elements of soil Physics

Koorevaar P, Menelik, G Dirksen C et al (1983) Elements of soil physics, 3rd edn. Elsevier, Amsterdam, 229 pp

Laane C, Boeren S et al (1987) Rules for optimization of biocatalysis in organic solvents. Biotechnol Bioeng 30(1):81–87

Lageman R, Pool W et al (1989) Electro-reclamation – theory and practice. Chem Indus (18): 585–590

Langenhoff AAM, Staps S et al (2002) Feasibility study of a “biological fence” at the site of shell Netherlands refinery – final report phase 2, extended monitoring SV-501. SKB

Lee ES, Seol Y et al (2003). Destruction efficiencies and dynamics of reaction fronts associated with the permanganate oxidation of trichloroethylene. Environ Sci Technol 37(11): 2540–2546

Liang CJ, Bruell CJ et al (2004a) Persulfate oxidation for in situ remediation of TCE. I. Activated by ferrous ion with and without a persulfate-thiosulfate redox couple. Chemosphere 55(9):1213–1223

Liang CJ, Bruell CJ et al (2004b) Persulfate oxidation for in situ remediation of TCE. II. Activated by chelated ferrous ion. Chemosphere 55(9):1225–1233

Mackay DM, Cherry JA (1989) Groundwater contamination – pump-and-treat remediation 2. Environ Sci Technol 23(6):630–636

MacKinnon LK, Thomson NR (2002) Laboratory-scale in situ chemical oxidation of a perchloroethylene pool using permanganate. J Contamin Hydrol 56(1–2):49–74

Malina G, Grotenhuis JTC et al (1998) Soil vapour extraction versus bioventing of toluene and decane in bench-scale soil columns. Environ Technol 19(10):977–991

Malina G, Grotenhuis JTC et al (2002) Vapor extraction/bioventing sequential treatment of soil contaminated with volatile and semi volatile hydrocarbon mixtures. Bioremediation J 6(2): 159–176

Marley MC, Hazebrouck DJ et al (1992) The application of in situ air sparging as an innovative soils and ground-water remediation technology. Ground Water Monitor Remed 12(2):137–145

McCarty PL (1991) Engineering concepts for in situ bioremediation. J Hazard Mater 28(1–2):1–11

McCarty PL, Semprini L et al (1991) In situ methanotrophic bioremediation for contaminated groundwater at St. Joseph, Michigan. In: Hinchee RE, Olfenbuttel RF (eds) On site bioreclamation; processes for xenobiotic and hydrocarbon treatment, Butterworth-Heinemann, Boston, MA, pp 16–40. ISBN 0-7506-9301-0

Middeldorp P, Langenhoff A et al (2002) In situ clean up technologies. In: Reineke W, Agathos S (eds) Focus on biotechnology, Series, vol 3a. Kluwer, Dordrecht. vi+241 pp. ISBN 1-4020-0529-6

Miller RN, Vogel CM et al (1991) A field scale investigation of petroleum hydrocarbon biodegradation in the vadoze zone enhanced by soil venting at Tyndall AFB, Florida. In: Hinchee RE, Olfenbuttel RF (eds) In situ bioremediation. Butterworth-Heinemann, Boston, pp 283–301

Norris et al (1994) Handbook of bioremediation. Lewis Publshers, Boca Raton, FL, 257pp. ISBN 1-56670-074-4 p, ISBN 1402061013; 9781402061011

Oldenhuis R, Oedzes JY et al (1991) Kinetics of chlorinated-hydrocarbon degradation by methylosinus-trichosporium Ob3b and toxicity of trichloroethylene. Appl Environ Microbiol 57(1):7–14

Pignatello JJ, Oliveros E et al (2006) Advanced oxidation processes for organic contaminant destruction based on the Fenton reaction and related chemistry. Crit Rev Environ Sci Technol 36(1):1–84

Plaisier W, Bennen M, Pancras T, Bouwhuis E (2003) Performing in-situ chemical oxidation with Fenton's reagent at a former gas work in Rotterdam. Conference proceedings ConSoil 2003, 8th international FZK/TNO conference on contaminated soil,pp 1700–1708

Ramakrishnan V, Ogram AV, Lindner AS (2005) Impacts of co-solvent flushing on microbial populations capable of degrading trichloroethylene. Envriron Health Perspect 113(1):55–61

Reid BJ, Stokes JD et al (2000) Nonexhaustive cyclodextrin-based extraction technique for the evaluation of PAH bioavailability. Environ Sci Technol 34(15):3174–3179

Roels JA, Kossen NWF (1978) On the modelling of microbial metabolism. Prog Ind Microbiol 14:98–203

Salanitro JP, Johnson PC et al (2000) Field scale demonstration of enhanced MTBE bioremediation through aquifer bioaugmentation and oxygenation. Environ Sci Technol 34(19):4152–4162

Scherer MM, Richter S et al (2000) Chemistry and microbiology of permeable reactive barriers for in situ groundwater clean up. Crit Rev Environ Sci Technol 30(3):363–411

Schlegel HG, Schmidt K (1985) Allgemeine Mikrobiologie. Thieme 571

Semple KT, Morriss AWJ et al (2003) Bioavailability of hydrophobic organic contaminants in soils: fundamental concepts and techniques for analysis. Eur J Soil Sci 54(4):809–818

Sinke A, Heimovaara T, Tonnaer H, Van Veen J (1998) Decision support system natural attenuation. CUR-NOBIS Report 97-1-02, The Netherlands

Smart JL (2005) Application of six-phase soil heating technology for groundwater remediation. Envriron Prog (1):34–43

Smit MPJ (2009) Contaminant release from sediments: a mass flux approach. PhD thesis, Wageningen University, Wageningen, The Netherlands. ISBN 978-90-8585-458-6

Smit MPJ, Grotenhuis T et al (2008) Desorption of dieldrin from field aged sediments: simulating flood events. J Soils Sediments 8(2):80–85

Starr RC, Cherry JA (1994) In-situ remediation of contaminated ground-water – the funnel-and-gate system. Ground Water 32(3):465–476

Ter Meer JTM, Van Duijne H, Nieuwenhuis R, Rijnaarts HHM (2008) Prevention and reduction of pollution of groundwater at contaminated megasites: integrated management strategy, and its application on megasite cases. In: Quevauviller Ph (ed) In groundwater science & policy: an international overview. RSC Publishing, Cambridge, the UK The Royal Society of Chemistry. pp 405–420

Van Agteren MH, Keuning S, Janssen DB (1998) Handbook on biodegradation and biological treatment of hazardous organic compounds. Kluwer, Dordrecht. ISBN 0-7923-4989-X

Van Heiningen WNM, Nipshagen AAM et al (1999) Intrinsic and enhanced biodegradation of benzene in strongly reduced aquifers. In: Alleman BC, Leeson A (eds) In situ bioremediation of petroleum hydrocarbon and other organic compounds. The fifth international in situ and on site bioremediation symposium, vol 5(3). Battelle, San Diego, pp 481–485

Van Leeuwen CJ, Vermeire TG (2007) Risk assessment of chemicals: an introduction, 2nd edn. Springer, Dordrecht, The Netherlands. ISBN 978-1-4020-6101-1 (handbook), ISBN 978-1-4020-6102-8 (e-book)

Van Nooten T, Diels L et al (2008) Design of a multifunctional permeable reactive barrier for the treatment of landfill leachate contamination: laboratory column evaluation. Environ Sci Technol 42(23):8890–8895

Van Wezel AP, Franken ROG, Drissen E, Versluijs CW, Van den Berg R (2007) Maatschappelijke Kosten-BatenAnalyse van de Nederlandse Bodemsaneringsoperatie. MNP Rapport 500122002/2007

Vangronsveld J, Colpaert JV et al (1996) Reclamation of a bare industrial area contaminated by non-ferrous metals: physicochemical and biological evaluation of the durability of soil treatment and revegetation. Environ Pollut 94(2):131–140

Vegter JJ (2004) Risk based land management, Status and perspectives for policy. In: Lens P, Grotenhuis T, Malina G, Tabak H (eds) Soil and sediment remediation. IWA Publishing, London, pp 313–330. ISBN1-84339-100-7

Vegter JJ, Lowe J, Kasamas H (eds) (2003) Sustainable management of contaminated land: an overview. Austrian Federal Environment Agency on Behalf of CLARINET, Vienna Austria. ISBN: 3-85457-674-9

Volkering F, Quist JJ et al (1998) A rapid method for predicting the residual concentration after biological treatment of oil-polluted soil. Contaminated Soil ’98, Thomas Telford, London, pp 251–259

Wilson JT, Ward CH (1986) Opportunities for bioremediation of aquifers contaminated with petroleum hydrocarbons. J Ind Microbiol 27:109–116

Internet Links

http://ec.europa.eu/environment/water/water-framework/index_en.html

http://www.epa.gov/glnpo/sediment/iscmain/

http://www.estcp.org/Technology/upload/ISO_Report.pdf 1999 Technology status review in situ oxidation

http://www.eurodemo.info/ (visited November 2009): Regularly updated information on path oriented In Situ technologies

http://www.vito.be/sedbarcah

http://www.irgc.org/IMG/pdf/IRGC_WP_No_1_Risk_Governance__reprinted_version_.pdf, p 37 and http://www.irgc.org/Publications.html

Chapter 22
Natural Attenuation

Anita Peter, Thomas Held, Norbert Hüsers, and Frank A. Swartjes

Abstract Natural Attenuation (NA) has emerged during the last 10–15 years as a useful and cost-efficient alternative approach for contaminated site management. It refers to the naturally occurring processes like dispersion, diffusion, sorption, volatilization, degradation and transformation, all of which can substantially decrease contaminant concentration, mass, toxicity and/or mobility within soil and groundwater. The efficiency of Natural Attenuation processes depends to a large extent on site-specific conditions, primarily on the type of contaminants present at the site. Proving and evaluating the efficiency of Natural Attenuation processes is a prerequisite for accepting them as the sole or additional remediation alternative. The implementation of Natural Attenuation as a remediation alternative, i.e. the monitoring that assures sustainability of Natural Attenuation processes over time, is called Monitored Natural Attenuation (MNA). This chapter presents an overview of the history and political acceptance of Natural Attenuation and the principles on which it was built. Specifically, it describes how the different processes act on contaminant plume development and explores methods of evaluating Natural Attenuation processes and proving their effectiveness. A stepwise approach to assess and implement Natural Attenuation is presented, followed by three sections on the most frequently found contaminant groups for which Monitored Natural Attenuation is being applied. These are petroleum hydrocarbons, chlorinated hydrocarbons and tar oil contaminants. Characteristics of these contaminant groups, the resulting contaminant-specific potential for implementation of Natural Attenuation and the challenges to be expected, are elucidated and discussed.

A. Peter (✉)
Institute for Geosciences, Christian-Albrechts-University, Ludewig-Meyn-Str. 10,
24118 Kiel, Germany
e-mail: anita.peter@gpi.uni-kiel.de

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_22,

Contents

22.1 Introduction

22.1.1 Principles

Natural Attenuation (NA) signifies the reduction of mass, toxicity, mobility, volume, and/or concentration of contaminants in soil or groundwater using naturally occurring processes in soil. These processes can be classified as *destructive* and *non-destructive*. Biodegradation, a process in which microorganisms break down contaminants, is the most well-known and prevalent destructive mechanism. These microorganisms use organic material, like organic contaminants, as a source for growth and metabolism (i.e., as a carbon and energy source). Also abiotic degradation processes like hydrolyses may occur. Non-destructive Natural Attenuation processes include diffusion, dispersion, and volatilization, resulting in dilution and spreading of contaminants, while sorption leads to a reversible or irreversible immobilisation of contaminants. By its nature, Natural Attenuation is rather a series of processes than a technology and its efficiency depends on the contaminants and on individual site-specific conditions.

Since early appraisals proved the power of natural degradation to contribute to soil and groundwater remediation, methods to stimulate these processes have been developed. Especially, technologies were tested and used to stimulate biodegradation, e.g., by increasing the number of soil and groundwater organisms (*bioaugmentation*) or more frequently by increasing the activity of organisms (*biostimulation*), i.e. by overcoming limiting factors as, for example, electron acceptors or donor shortage. Today, this *in situ* remediation technology is also known as '*Enhanced Natural Attenuation' (ENA)*. It is generally acknowledged that, in cases where the efficiency of *Natural Attenuation* has been shown, monitoring is necessary to demonstrate that Natural Attenuation works in a sustainable manner. Therefore, the term *Monitored Natural Attenuation* (*MNA*) is used.

Monitored Natural Attenuation is primarily applied at sites that have been contaminated by readily degradable contaminants, i.e. organic contaminants such as volatile organic contaminants (VOCs), semi-volatile organic contaminants (sVOCs), including especially fuel hydrocarbons such as gasoline, jet fuel or diesel fuel. To a lesser extent it has even been used for less degradable contaminants such as chlorinated hydrocarbons, methyl-tert-butylether (MTBE), polychlorinated biphenyls (PCBs) and explosives (Pennington et al. 1999). But also for metal and metalloids, permanent immobilization within a biogeochemical active zone might be considered as an attenuation process. The observed immobilization is caused by a change in the valence state of the metal/metalloid. The change in the valence state may not only reduce the solubility and therefore the mobility of the inorganic contaminant, but also its toxicity. Chromium, for example, is reduced in anaerobic environments from the highly mobile and toxic Cr(VI) form to the low soluble and low toxic Cr(III) form (Laitine 2006).

22.1.2 History

In fact, before Monitored Natural Attenuation emerged as an alternative to active remediation measures, Natural Attenuation processes had always taken place. From the moment that contaminants enter soil and groundwater, they are spread and diluted into the environment, immobilised and degraded by microorganisms. Already in the early half of the last century it was recognized that organic contaminants are degraded in groundwater. The realization that anthropogenic contaminants may enter and pollute the groundwater started in the 1970s. With regard to contaminated site management, the possibilities of NA were also recognized at an early stage. In the 1980s, several experiments that focused on the degradation of organic contaminants by organisms were performed in North America and Europe. This new phenomenon was illustrated, for example, by Wilson and Wilson (1985), who demonstrated the aerobic degradation of trichloroethene in an unsaturated soil column, while in Europe, e.g., Valo et al. (1985) studied the requirements and conditions for pentachlorophenol (PCP) biodegradation in Finish soils by a mixed bacterial culture. Since organisms are ubiquitous and work for free, most of these experiments were triggered by the cost efficiency aspect. Since the mid 1980s, and certainly since the mid 1990s, regulators and environmental engineers are increasingly relying upon MNA as an alternative 'remediation' approach. This ended in the US in the so-called OSWER directive (office of solid waste and emergency response) published by the US EPA in 1999, where the usage of MNA was regulated. At the same time several technical protocols were published (AFCEE 2000; Wiedemeier and Chapelle 1998; Wiedemeier et al. 1995; US EPA 1998). In Europe, the idea of Natural Attenuation was taken up very early in various countries (Rügner et al. 2006). In Germany, a huge research and development program was launched in 2002 by the Ministry for Education and Research (BMBF) on the technical and legal application of MNA in Germany. As a result, seven contaminant-specific technical guidelines as well as a recommended course of action were produced (see www.natural-attenuation.de).

22.1.3 Definition

Many institutions have created their own definitions of Natural Attenuation. Most of them came up with nearly identical definitions. Basically, Natural Attenuation is an approach that relies on natural processes to attenuate contamination in soil and groundwater. The U.S. Environmental Protection Agency (OSWER) defines MNA as the 'reliance on Natural Attenuation processes (within the context of a carefully controlled and monitored site cleanup approach) to achieve site-specific remediation objectives within a time frame that is reasonable compared to that offered by other more active methods' (US EPA 1999). According to this definition the NA processes include in situ occurring physical, chemical, or biological processes, e.g., biodegradation, dispersion, dilution, sorption, volatilization, radioactive decay and chemical or biological stabilisation, transformation, or destruction of contaminants. Under favourable conditions these processes act without human intervention.

The American Society for Testing and Materials (ASTM) defines Natural Attenuation as the 'reduction in mass or concentration of a contaminant in groundwater over time or distance from the source of constituents of concern due to naturally occurring physical, chemical, and biological processes, such as; biodegradation, dispersion, dilution, adsorption, and volatilization' (ASTM 2004). The ASTM-definition did not contain guidance to the time MNA should be applied.

The U.S. Air Force Centre for Environmental Excellence defines NA as the processes resulting 'from the integration of several subsurface attenuation mechanisms that are classified as either destructive or non-destructive. Biodegradation is the most important destructive attenuation mechanism. Non-destructive attenuation mechanisms include sorption, dispersion, dilution from recharge, and volatilization' (Wiedemeier et al. 1999).

In Germany, the official wording is based on the OSWER directive (LABO 2009). Natural Attenuation processes are physical, chemical and biological processes which lead – without human intervention – to a reduction of mass, toxicity, mobility, volume or concentration of a contaminant in soil and groundwater.

22.1.4 Political and Practical Acceptance

At present, Monitored Natural Attenuation is widely recognized and used as an alternative to or in combination with active technologies. Especially, the cost-effectiveness of MNA is appreciated at the first glance. However, in some countries (e.g. Germany) comprehensive site investigations exceeding those for typical active remediation measures are required to fulfil regulatory demands (see also Section 22.2.2). Together with a well-founded prediction by reactive transport modelling and a long lasting monitoring program, MNA might become as expensive as active remediation measures.

Despite this uncertainty concerning the costs, using natural processes for contaminated site management has some additional serious drawbacks. Firstly, the release of contaminants out of a NAPL source into soil and groundwater due to dissolution is a long lasting process under natural conditions. Time scales for NAPL dissolution range from tens of years for residual NAPL blobs to hundreds to thousands of years for NAPL pools (see Grathwohl 1998). When released to groundwater, the degradation of contaminants is usually very fast, as long as reaction partners (i.e. electron acceptors or donors) are present. After consumption of reaction partners the delivery of additional reaction partners occurs via diffusion across the fringes of the contaminant plume, which is a slow process. As a consequence of the long lasting contaminant release and the slow delivery of reaction partners that act as a limiting factor for biodegradation, contaminated sites (more precisely: the groundwater) are often not suitable for their purpose for many decades. During this period, control measures might be necessary to ensure that no unexpected risks for human health and the ecosystem may arise. Another major problem is the fact that the time span at which natural processes achieve the maximum cleanup level (MCL) is very difficult to predict as it is directly related to the source lifetime. Both the physico-chemical processes which are responsible for the contaminant release and transport and the biological degradation processes are difficult to quantify. Moreover, the extent to which processes such as competing degradation of different contaminants or bioavailability of different forms of electron acceptors in the mineral phase (e.g. FeIII or MnIV) influence plume development, are often largely unknown.

Another problem that may arise when implementing MNA is that some degradation interim metabolites are more harmful than the original contaminants, e.g., the carcinogenic and mobile vinyl chloride as a degradation metabolite of chlorinated solvent such as perchloroethene (PCE). As a consequence, MNA can only be accepted in case the degradation does not end with the accumulation of metabolites.

An interesting political dilemma is related to *dilution* of contaminants, caused by migration of contaminants in soil and groundwater. On the one hand, dilution results in lower concentrations, which is often the major goal of Risk Management. On the other hand, larger volumes of groundwater become contaminated through dilution. To overcome this contradiction, besides reduction of concentrations often a reduction in mass flow rates is required in some countries before accepting MNA. This can only be achieved when other processes, i.e. degradation, volatilization and sorption, also contribute to the concentration reduction. Hence, substantial importance should be given to these processes and their quantification.

The circumstances under which MNA is accepted as the sole, rather than an accompanying, measure for contaminant site management varies for different countries. The US EPA (1999) emphases that 'MNA should not be considered a default or presumptive remedy at any contaminated site'. Source removal or control techniques remain an important measure to address the principal threat where applicable, and MNA is only appropriate when the remediation timeframe is reasonable compared to other alternatives.

In Germany, the time scale for MNA alone is not decisive, however the proportionality of active remediation measures becomes decisive regarding decisions on the implementation of MNA. This means, that MNA can only be accepted as a stand-alone measure if other active remediation technologies are not proportional, i.e. if they are too expensive in relation to their efficiency (LABO 2009).

In the Netherlands Natural Attenuation falls within the responsibility of the regional authorities. An important boundary condition based on national legislation, however, is that unacceptable human health risks must be excluded and the consequences of Natural Attenuation must be monitored.

22.2 Principles of Natural Attenuation

22.2.1 Plume Development and Transport Processes

Release of contaminants into soil and/or groundwater usually leads to formation of contaminant plumes in groundwater. A contaminant plume is defined here as the volume of groundwater containing a non-zero quantity of the contaminant. This plume shows different states of development, i.e., an expanding, a steady-state or a shrinking phase (see Fig. 22.1).

These phases of plume development depend on the mass balance of contaminants for the groundwater compartment (Fig. 22.2). The plume is in steady state when the input rate of contaminants into groundwater equals the output rate of contaminants from groundwater (individual rates discussed below). If the input rate exceeds the output rate the plume will extend, while a higher output than input rate will cause a shrinking plume.

Inputs (into groundwater) result from dissolution from NAPL, desorption or a direct release of dissolved contaminants via groundwater recharge or leachate. Outputs (from the groundwater) are the above mentioned Natural Attenuation processes (see Section 22.1.1), however (in a strict sense) excluding dispersion and diffusion (Fig. 22.2). While dispersion and diffusion lead to a spreading of contaminants and thus to a reduction of contaminant concentrations, they do not lead to a

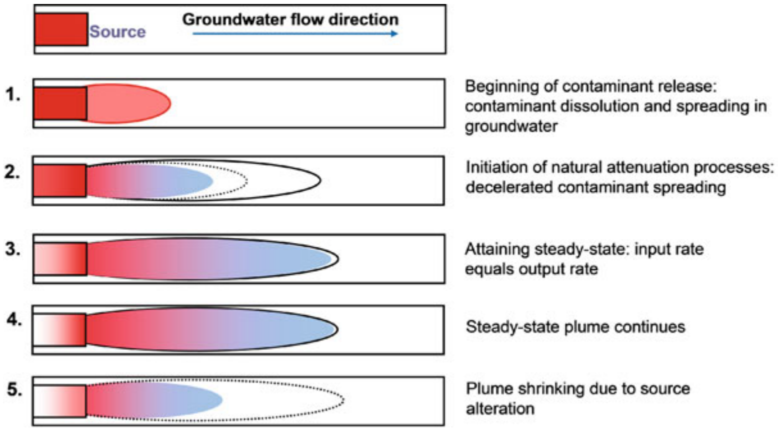


Fig. 22.1 Temporal development of a contaminant plume in groundwater due to Natural Attenuation processes. Plume expansion occurs during time steps 1 and 2, during time steps 3 and 4 the contaminant plume has reached steady-state conditions, while the plume shrinks in step 5 (adapted from Teutsch and Rügner (1999))

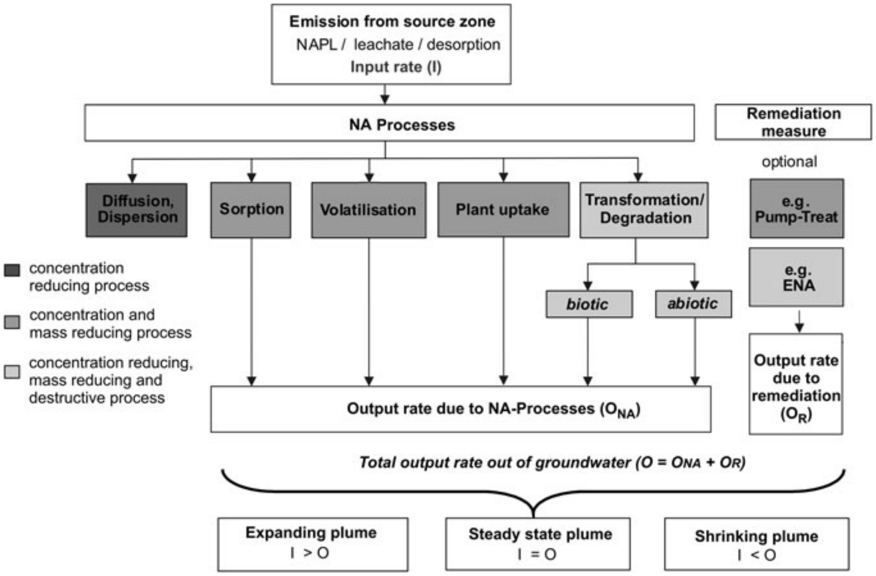


Fig. 22.2 Overview over contaminant input and output processes and their role with respect to contaminant plume development. Active remediation measures contribute optionally to a contaminant output in groundwater, either by contaminant destruction (e.g., ENA) or by shifting contaminants into other compartments (e.g., Pump and Treat) (adapted from Grandel and Dahmke (2008))

reduction in contaminant mass within the groundwater, hence they do not contribute to the overall rate of output from groundwater. But in the broader sense, when defining a contaminant plume as the volume of groundwater that contains contaminants above a certain limit, e.g., the detection limit of the contaminant, the plume's outline shown by this concentration isoline can become stationary just due to dispersion and diffusion processes. An analytical solution to calculate the time that is needed to achieve steady-state conditions for a certain concentration isoline only considering dispersion and diffusion can be found in (Domenico 1987).

Sorption, volatilisation or uptake by plants entail not only a contaminant concentration reduction, but also a contaminant mass reduction in the groundwater by shifting the contaminants into another compartment (i.e., into sediment, soil gas or plants). The sole destructive processes that reduce contaminant mass, not only by shifting them into other compartments, but by degrading them into – hopefully – harmless end products, are biodegradation and chemical transformation.

The way Natural Attenuation processes act on the development of a contaminant plume in groundwater is schematically depicted in Fig. 22.3. A mathematical description of the transport processes and explanation of important parameters influencing these processes is found in Rolle et al. (Chapter 19 of this book).

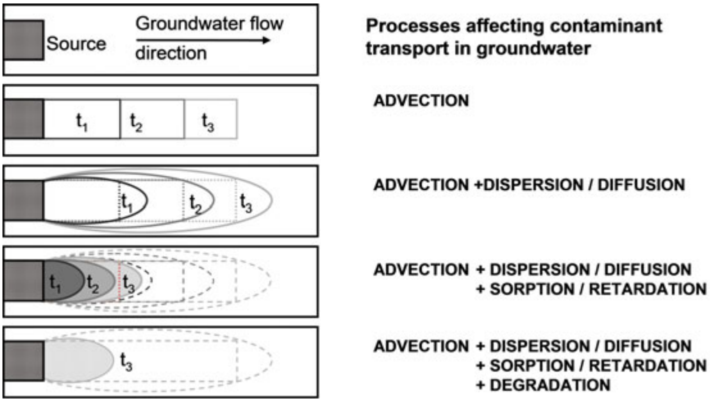


Fig. 22.3 Schematic illustration of the impact of transport processes on contaminant plume development in groundwater at different time steps (Teutsch et al. 1997)

22.2.2 *Proving Natural Attenuation and Implementing Monitored Natural Attenuation*

Considering Natural Attenuation as a sole or additional remediation option, two key questions concerning the source zone and the evolving contaminant plume have to be addressed:

- How far will the contaminant plume migrate from the source zone, i.e., which receptors are likely to be affected?
- How long will the plume persist, until a certain acceptable 'clean-up' level is achieved?

Answers to these key questions will require answers to many sub-questions, which are all highly site-specific.

According to the US EPA (1999) several indicators or lines of evidence have been identified that can be used to prove that natural processes are effective. At least the first two of the following lines of evidence are normally required to document NA efficiency (U.S. EPA 2009):

- Historical trends indicate a decrease in contaminant concentrations and/or mass over time and revealing a plume whose spatial extension is stable or shrinking over time. A stable or shrinking plume indicates that biodegradation and other mass flux reducing processes, e.g. volatilization and sorption, are removing dissolved contaminants from the groundwater at an equal or greater rate than the source is adding them to the plume.
- Contaminant concentration and mass decrease with distance as well as chemical indicators: Biodegradation of contaminants is directly related to changes in groundwater chemistry such as the biological consumption of natural levels of electron acceptors (like dissolved oxygen, nitrate, manganese (IV), iron (III), and sulphate) and the formation of by-products such as dissolved iron (II), manganese (II), and methane. These geochemical redox indicators can be used to estimate the site-specific potential for contaminants to be mineralized by biodegradation.
- Laboratory microcosm studies. These studies can be used to investigate aquifer conditions under more controlled conditions and to demonstrate that native bacteria can biodegrade contaminants of concern. In microcosms, degradation rates can be measured, allowing a comparison of relative degradability for different contaminants. Generally, this technique is recommended when one of the first two lines of evidence is inconclusive.

Besides this 'Three Lines of Evidence'-approach several additional NA protocols exist to address the above mentioned questions (NRC 2000; Rice et al. 1995; USEPA 1994, 1999, 2005; Wiedemeier and Chapelle 1998; Wiedemeier et al. 1995). In the following, the comprehensive protocol developed within KORA, a six year priority research program on NA and MNA (http://www.natural-attenuation.de) funded by the German Ministry of Education and Research, is briefly presented. More detail is found in (Michels et al. 2008) and (Wabbels and Teutsch 2008). This approach for investigating NA and implementing MNA is divided into the following main steps:

I. Checking the prerequisites for MNA.
II. Site investigations to prove the effectiveness of NA.
III. Prediction and decision concerning MNA.
IV. Monitoring and site closure.

I. *Checking the prerequisites for MNA* implies the answers to the following questions:

   (1) Are the general conditions concerning actual and future use of the site and surroundings including its groundwater suitable to implement MNA?
   (2) Are the existing contaminants degradable, in principal?

   Furthermore, a first spatial delineation of the NA reaction zone, i.e., the source zone and the contaminant plume, has to be performed, including a rough characterization of the geochemical and hydraulic situation. Also, a first conceptual site model including the hydraulic and geochemical data has to be developed. This first step results in a primary evaluation of the NA potential at the site.

II. *Site investigations to prove the effectiveness of NA*. Aims of this step are:

   (1) Identification and quantification of NA processes.
   (2) Proof of decreasing contaminant mass flow rates.
   (3) Generation of data basis for NA prediction.

   This implies the application of various methods, which depend in part on the site and/or contaminants present. Table 22.1 and Section 22.2.3 summarize methods that have been proven suitable at various test sites (Wabbels and Teutsch 2008, Peter et al. 2006).

III. *Prediction and decision concerning MNA* include the following steps:

   (1) Set up of a flow and reactive transport model to simulate future source and plume development.
   (2) Proof of proportionality: only if active remediation options are not proportional, i.e., if they are too expensive in relation to their efficiency, MNA might become a remediation alternative. Thus, active remediation options have to be considered and compared to MNA.
   (3) Design and implementation of a MNA concept with definition of targets and fall-back criteria.

IV. *Monitoring and site closure* comprise the following aspect:

   (1) Monitoring with data assembly for conservation of evidence.
   This target performance comparison can eventually lead to an end of monitoring and a site closure, or to an abortion of MNA and a return to active remediation measures.

22.2.3 Methods to Prove Monitored Natural Attenuation

As mentioned above in Section 22.2.2 several lines of evidence may lead to the proof of Natural Attenuation, entailing a wide portfolio of site and laboratory investigations, often accompanied by modelling. In Table 22.1 investigation methods are summarized that have been shown suitable at several sites contaminated with petroleum hydrocarbons or other contaminants to identify and quantify Natural

Table 22.1 Summary of investigation methods to prove the effectiveness of NA at petroleum hydrocarbon contaminated sites

Aim		Method	Comments
Source zone characterization	Source localization	Investigation of soil and groundwater samples	Standard method, invasive and destructive method, direct measurements
		Direct-push (DP) investigations using in-situ contaminant detectors (LIF / ROST or MIP-PID, -FID, -DELCD)	Invasive, but non-destructive method, indirect measurements, influenced by soil and contaminant characteristics
		In-situ radon analyses in the soil gas to detect radon depletions due to NAPL presence	Indirect measurement of NAPL presence, depend on soil characteristics
	NAPL mobility	Pumping tests w/o tracer tests in the NAPL phase	Only in highly permeable aquifers with significant NAPL thicknesses suitable
	Source mass	Calculation of mass from 3-phase profiles using apparent NAPL phase thicknesses from groundwater monitoring wells and pedo-physical parameters	Indirect calculation, based on measured parameters
	NAPL composition	NAPL phase analyses and derivation of contaminant specific saturation concentrations	Standard method
		Column experiments to measure saturation concentrations	For confirmation of NAPL phase analyses
	Source emission	Tank experiments to measure source emission	Highly laborious method with restricted transferability to the field
		Analytical and/or numerical calculation of source emission from NAPL	Indirect calculation based on various measured parameters; numerical calculations highly complex
Plume characterization	Spatial and temporal plume delineation	Conventional groundwater sampling	Standard method, good reproducibility
		Multilevel groundwater sampling using e.g., DP low cost and small diameter monitoring wells	Often special pumping equipment necessary

Table 22.1 (continued)

Aim		Method	Comments
		DP groundwater sampling using, for example, BAT groundwater sampler	Not suitable for long-term plume monitoring as measurement is singular
		In-situ semi-quantitative groundwater analyses using DP-MIP	Not suitable for long-term plume monitoring as measurement is singular
	Mass flow rate reduction	Integral pumping tests at two or more consecutive control planes with increasing distance to the source zone	Integral pumping reduces uncertainty of mass flow rate determination due to spatial integration
		Passive samplers in multiple depths or multilevel groundwater sampling in conjunction with depth profiles of hydraulic conductivity, for example, by DP slug tests or DP injection logging	Temporal integration of concentration by passive samplers leads to average temporal concentrations and thus average mass flow rates
NA process identification and quantification	Proof of plume stationarity	Time series analyses (e.g., Mann-Whitney-Test or Mann-Kendall-Test)	Nonparametric statistical tests applied to concentration time series to prove plume stability. At least 4 sampling events are needed.
	Qualitative NA proof	Distribution of electron acceptors (depletion within source and plume) and reduced species of electron acceptors (accumulation within source and plume)	Standard method
		Comparison between reactive and conservative plume development	If a (measured or calculated) conservative plume is longer than the contaminant plume(s), this is a hint for active NA processes
		Presence of metabolites	Hint for active NA processes
		Chemical fingerprinting (spatial distribution of ratios of certain contaminant concentrations)	Shift of ratios indicate the preferred attenuation of certain contaminants

Table 22.1 (continued)

Aim		Method	Comments
		Microbiological analyses (cell population analyses, microcosm studies, in-situ respiration tests, Fluorescence in-situ hybridisation (FISH))	Results from laboratory studies are not directly transferable to evaluate degradation in the field
	Quantitative NA proof	Mass flow rate reduction (see above)	
		Determination of biodegradation rates from field data	Knowledge of flow paths required
		Quantification of total oxidation capacity using consumed electron acceptors and produced reduced species of electron acceptors	Spatial (2D or 3D) interpolation of electron acceptors necessary
		Quantification of percentage of biodegradation from contaminant specific isotope analyses	Direct, quantitative proof of degradation (only applicable to molecules up to C_{10}). Contaminant specific fractionation factors required (if no fraction factors are known, a qualitative evaluation of contaminant specific degradation is possible).
		Quantification of sorption processes (*kd* concept / non-linear sorption)	Sorption is highly sensitive on the type of organic material
	Reactive Transport Model	Quantification of coupled transport processes on the basis of measured distributed parameters	A lot of model codes with different levels of complexity available. Only method to consider all transport processes quantitatively and coupled.

DP Direct Push, *LIF* Laser Induced Fluorescence Probe, *ROST* Rapid Optical Screening Tool, *NAPL* Non-Aqueous Phase Liquid, *MIP* Membrane Interface Probe, *PID* Photo Ionization Detector, *FID* Flame Ionization Detector, *DELCD* Dry Electrolytic Conductivity Detector

For detailed descriptions and references see Wabbels and Teutsch (2008), Michels et al. (2008), Peter et al. (2006), Held (2007) and Werner et al. (2008a)

Attenuation processes. These methods address the topics as outlined above in Section 22.2.2 to first delineate the contaminant source zone and plume development and then to show qualitatively and quantitatively that Natural Attenuation processes take place.

Most important qualitative hints on active NA processes are the first two out of the 'Three Lines of Evidence' (Wiedemeier et al. 1995), i.e., the proof of plume stability or shrinkage over time and the contaminant and reaction partner reduction versus distance accompanied by increasing daughter contaminant concentrations.

As quantitative methods to prove NA processes, several NA protocols call for a proof of contaminant mass flow rate reduction at several control planes consecutively down-gradient of the source. This can be realized either by the so called 'groundwater fence' method (e.g., King et al. 1999) or by integral pumping tests (e.g., Bockelmann et al. 2000).

Further quantitative methods to obtain evidence on degradation activities are the determination of degradation rates from field data, where applicable by using a tracer correction term according to (Wiedemeier et al. 1999), or the contaminant-specific isotope analyses to quantify via contaminant fractionation the share of degradation that has occurred (e.g., Hunkeler et al. 2001).

The sole possibility to quantify all NA processes and their interactions and to predict future plume development is by employing reactive transport modelling. Various reactive transport models exist that have incorporated NA processes at different levels of detail. An example for comprehensive reactive transport modelling considering various degradation reactions can be found, for example, in Miles et al. 2008.

22.3 Natural Attenuation at Petroleum Hydrocarbon Contaminated Sites

22.3.1 Characteristics of Petroleum Hydrocarbon Mixtures

Distillation of crude oil leads to various petroleum hydrocarbons products, exhibiting a large range of different contaminant compositions and thus different physico-chemical behaviour. Gasoline is the lightest distillate (40–175°C), followed by kerosene (150–280°C), Diesel (160–390°C) and lube oil (300–525°C). Petroleum hydrocarbon products consist mainly of aliphatics, which can be divided into alkanes, alkenes and cyclo-alkanes, and aromatics, i.e., mono- and polycyclic aromatics. Aliphatics constitute with 40–50 weight percentage the main contaminant group, whereas aromatics have shares of 22–35 weight percentage (Potter and Simmons 1998). Additionally, additives and blending agents are added to petroleum fuel products to improve their desired characteristics. Due to its environmental relevance, the best known additive is methyl-tert-butylether (MTBE), used as anti-knock agent in gasoline.

Petroleum hydrocarbons have densities ranging from 0.7 g/ml (gasoline) to 0.9 g/ml (lube oil), i.e., they are lighter than water. Due to their low solubility in water, petroleum hydrocarbons constitute a separate phase in groundwater, termed lighter than water non-aqueous phase liquids (LNAPL).

After being released into the environment, petroleum hydrocarbons will be subject to weathering processes, i.e., dissolution in water, volatilization and biodegradation. Each process affects the component groups differently. As alkanes are less soluble but more volatile than aromatics of comparable molecular weight, alkanes will be found primarily in the gas phase, while aromatics constitute the main contaminants in the water phase. Generally, solubility and volatility decrease with increasing number of carbon atoms, i.e., with increasing molecular weight. Thus, as the more volatile and more soluble constituents disappear first from the petroleum mixture, a shift towards heavier constituents in the remaining NAPL occurs. The rate of dissolution and volatilization of a certain component is, according to Raoult's law, proportional to its solubility, respectively vapour pressure and its mole fraction within the NAPL mixture (see Chapter 19 by Rolle et al., this book). Gasoline contains for example only about 1% of benzene, which results in a saturation concentration of benzene in water of about 20 mg/L, due to its high maximum water solubility of 1,800 mg/L.

After dissolution and/or volatilization from the NAPL phase, the single constituents are subject to biodegradation especially in the water and/or to sorption onto the solid phase, i.e., the sediments. Most petroleum hydrocarbon components apparent in the water phase are readily biodegradable, especially under aerobic conditions. Biodegradability of the different petroleum hydrocarbon constituents decrease in the following order: straight-chained aliphatics, branched aliphatics, aromatics, cyclic aliphatics, MTBE. Anaerobic conditions occur after oxygen is depleted, leading to a thermodynamically steered order of electron acceptor consumption: if nitrate is present, it will be the next favourable electron acceptor after oxygen, before iron(III)/manganese(IV) and sulphate. After complete depletion of these electron acceptors, methanogenesis will eventually take place, where the substrate acts as both electron acceptor and donor. It is thermodynamically the least favourable degradation pathway, exhibiting thus the slowest degradation kinetics.

Under anaerobic conditions MTBE, benzene and trimethylbenzenes are less degradable, which may cause long groundwater contaminant plumes. Typical plume lengths of most petroleum hydrocarbon constituents are less than 300 m (Rice et al. 1995; Teutsch et al. 1997)), except those of MTBE, showing plume lengths of up to 1,800 m (Stupp 2007).

Sorption of petroleum hydrocarbons takes place on the organic material in the sediment. Thus, the adsorbed amount of a single constituent depends on the content of organic material in the sediment as well as on the contaminant's partitioning coefficient between water and the organic material (k_{oc}). Several empirical correlations exist to derive the k_{oc} coefficient from the better measurable octanol-water partition coefficient k_{ow} (Grathwohl 1998; see Chapter 19 by Rolle et al., this book). Generally, the partitioning coefficient k_{ow} increases with increasing molecular weight, thus polycyclic aromatic hydrocarbons tend to adsorb much more strongly than mono-aromatics or MTBE.

22.3.2 Natural Attenuation Potential and Challenges at Petroleum Hydrocarbon Contaminated Sites

Besides the specific experience gained using a variety of methods to investigate the efficiency of NA at petroleum hydrocarbon contaminated sites (see Table 22.1), some general conclusions concerning the NA potential and its limits can be drawn for petroleum hydrocarbon contaminated sites:

Concerning the s*ource zone*, the following experiences have been made at various test sites:

- At most petroleum hydrocarbon contaminated sites LNAPLs are present, which act as a long-lasting source for contaminants in groundwater and soil air. However, once dissolved in groundwater, most petroleum hydrocarbon constituents are readily biodegradable.
- Unlike DNAPLs, where it is barely possible to investigate the source zone architecture accurately in the field, several methods exist to delineate the lateral spreading of an LNAPL (see Table 22.1). However, assessing the LNAPL mass in the subsurface remains a challenge.
- Emission, i.e. the mass flow rate from an LNAPL source into groundwater or soil gas, from a typical petroleum hydrocarbon LNAPL differs for each component and is time-dependent, among other things due to temporally varying mole fractions within the LNAPL. Thus, mole fractions of relatively low soluble components (e.g., trimethylbenzenes) increase with time due to dissolution and depletion of more soluble components (e.g., BTEX), thus emission of the first mentioned constituents may increase with time for a certain time period. Generally, alkanes constitute the main contaminants in the soil air, while aromatics are the main groundwater contaminants. Besides BTEX and PAH, trimethylbenzenes are important contaminants in groundwater, which should be analysed at each petroleum hydrocarbon contaminated site.
- Emission from a petroleum hydrocarbon LNAPL is difficult to measure in the field, therefore alternative methods like analytical or numerical calculations and/or laboratory experiments have to be applied. The emission from an LNAPL source depends on parameters that characterize the unsaturated and the saturated zone as well as the NAPL. As an LNAPL in and below the capillary fringe constitutes a multiphase system (water-NAPL-air), parameters to mathematically describe this multiphase system are needed: these are VanGenuchten or Brooks Corey parameters for the phase saturation – pressure relationship and the saturation – relative conductivity relationship, as well as scaling factors depending on the surface tensions of the present fluids. Furthermore, groundwater flow velocities, porosities and recharge rates are needed, that all are highly site-specific (see, for example, Peter et al. 2008).

- Emission is almost independent of the LNAPL mass, but depends on the source zone extension. Thus a reduced LNAPL saturation, e.g. as a result of phase extraction, will not in the short term lead to significantly reduced mass flow rates out of the source zone, neither to reduced concentrations in the plume nor a reduction in plume length. This is because saturation concentrations of a specific contaminant in the groundwater as a result of NAPL dissolution depend on the compound's molar fraction and its solubility, and not on NAPL saturation. If, however a source zone remediation measure achieves not only a reduced LNAPL saturation, but also a reduction in source zone extension (e.g. partial source removal by Dig and Dump), then a reduced mass flow rate out of the remaining LNAPL will occur. In any case, whether the partial source removal leads only to decreased NAPL saturation or indeed to a reduced source zone extent, a reduced LNAPL mass translates directly into a shortened source lifetime and consequently also shortened plume lifetime.

With respect to the contaminant *plume in the groundwater*, the following general conclusions can be drawn:

- Generally, petroleum hydrocarbons are readily degradable, i.e. sites contaminated by petroleum hydrocarbons generally may come into consideration for MNA. However, also less degradable contaminants such as MTBE or trimethylbenzenes, forming long plumes and possibly hazardous metabolites, might occur within petroleum hydrocarbon mixtures, posing a challenge for MNA at those sites.
- For plume delineation, a high number of data with some uncertainty is better than a small number of highly precise data, improving interpolation or calibration results. Thus, low-cost methods to investigate the plume, such as Direct-Push methods, are recommended for a first detailed plume investigation.
- Although plume stationarity is one of the essential prerequisites for accepting and implementing MNA, this stationarity is hard to prove. Meta studies have shown that steady state plumes seem to be prevailing at about one third of 271 investigated petroleum hydrocarbon contaminated sites (Rice et al. 1995). However, most processes that influence plume development seem to be transient, e.g., source emission (see above), hydraulic conditions (e.g., changing groundwater flow velocities and/or directions, groundwater recharge), biodegradation due to changing geochemical conditions (e.g., due to depletion of immobile electron acceptors like Fe(III) or Mn(IV)), sorption due to the kinetic nature of sorption processes and/or due to backfilling of sorption places and potentially subsequent desorption. Transience of these processes occurs at time scales that might be shorter than the monitoring frequency (e.g., changing hydraulic conditions) or longer than the monitoring frequency (e.g., changing source emission, biodegradation, sorption). Thus, it can be expected that a transient plume development might not be detected due to too long monitoring frequencies or due to too short monitoring periods relative to the expected time scale of process instationarity.

- Considering all processes influencing source and plume behaviour (NAPL dissolution, hydraulic conditions, biodegradation, sorption, volatilization etc.) leads to very complex reactive multi-component transport models with very long computation times (e.g., Miles et al. 2008). Thus, a simplification using only the dominant processes at a site is necessary. However, simplifications are sensitive to the simulated plume development and thus plume prediction, so that they should be done carefully and qualified.

As outlined above, the main challenges for implementing MNA at petroleum hydrocarbon contaminated sites are the proof of plume stationarity as well as a prediction of the NA processes, i.e., of plume development. This is especially true when being aware of usually short monitoring time periods (in the order of several years) that are used to calibrate transport models to predict time periods in the order of tens to perhaps hundreds of years.

22.4 Natural Attenuation at Chlorinated Hydrocarbon Contaminated Sites

22.4.1 Characteristics of Chlorinated Hydrocarbons

After being released into the subsurface, chlorinated volatile organic contaminants (CVOCs) will migrate as dense non-aqueous phase liquids (DNAPL) through the aquifer and – depending on the released volume – down to the aquitard where they spread and possibly form free product pools according to the aquitard morphology. These pools serve as a long lasting reservoir for CVOC contamination of the aquifer as dissolution of DNAPL pools occurs within the order of tens to thousands of years (Eberhardt and Grathwohl 2002). Solubilization of residual CVOC phase droplets and ganglia left in the aquifer is much faster (within the order of years to few tens of years), however due to their vertical geometry the dissolution of contaminants may result in a contaminant plume distributed over the entire thickness of the aquifer.

When dissolved in water, chlorinated hydrocarbons are subject to reductive dechlorination under reducing conditions (Bradley 2003) in a way that successively one chlorine-atom by another is replaced by a hydrogen atom. In case tetrachloroethene (PCE) represents the parental contaminant, the reductive dechlorination leads via trichloroethene (TCE), dichlorothene (DCE) and vinylchloride (VC) to the formation of ethene and (under extremely reducing conditions) also ethane (Fig. 22.4). Biological degradation of PCE or TCE predominantly leads to the formation of the *cis*-1,2-dichloroethene (cDCE) isomer, whereas the trans-isomer is formed to a much smaller extent only. The last steps in the reductive degradation pathway – i.e. degradation of DCE and VC – require strong reducing conditions (sulphate reduction/methanogenesis). Unlike the petroleum hydrocarbons, CVOCs do not serve as electron donors in this degradation pathway, but as electron acceptors; hence this biodegradation reaction is called dehalorespiration. The hydrogen

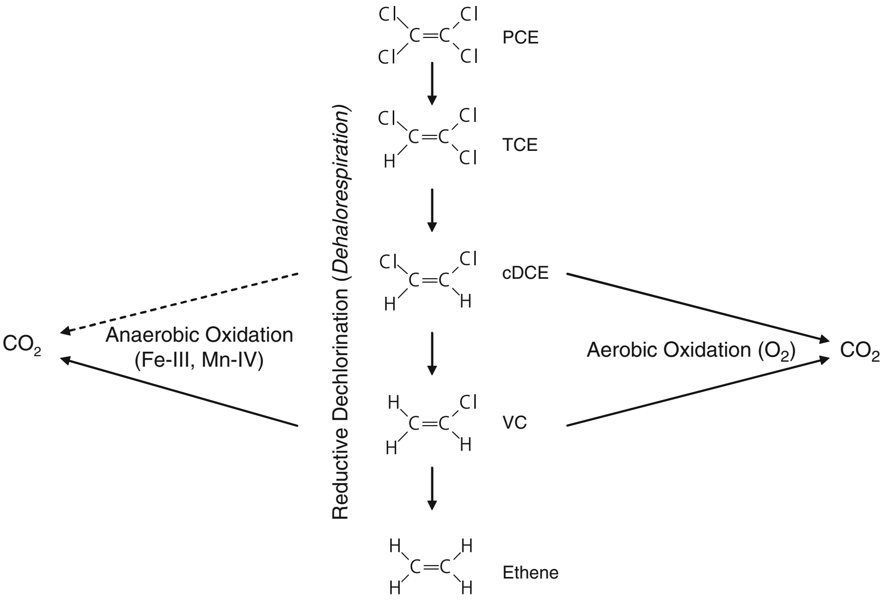


Fig. 22.4 Schematic overview of biological degradation pathways of chlorinated ethenes

required for the degradation is supplied during the degradation of organic non-chlorinated contaminants originating from natural organic matter (NOM) or from additional anthropogenic contaminations, like e.g. BTEX.

Apart from this major degradation pathway, low-chlorinated contaminants like cDCE and VC may also be degraded under aerobic conditions (Hanson and Brusseau 1994), either co-metabolically with e.g. methane as primary substrate (inducer) or productively (i.e. these substances serve as carbon and energy source) which has been observed in the lab and field at least for VC (e.g. Davis and Carpenter 1990). Furthermore, cDCE and/or VC seem to be subject to anaerobic oxidative degradation to CO_2 in various biogeochemical environments (manganese reduction, iron reduction, sulphate reduction) (Bradley and Chapelle 1996, 1997; Hata et al. 2003) (Fig. 22.4). It should be noted that this degradation pathway has been observed so far only in lab experiments. Hence, the relevance of anaerobic oxidative CVOC degradation for MNA is considered to be small, as well as the aerobic co-metabolic (but not productive) CVOC degradation.

Considering suitable conditions for biodegradation of CVOCs, the geochemical situation at CVOC contaminated sites often shows a redox environment which changes to more oxidizing conditions with increasing distance from the source due to impoverishment of electron donors originating from point sources. However, one crucial requirement – besides the presence of microorganisms that are capable to degrade CVOCs – for complete anaerobic CVOC degradation is the opposite: with increasing distance from the source the portion of low chlorinated metabolites increases more and more and their reductive dechlorination to ethene and ethane requires an increasingly reductive redox potential. This might be an important reason why, in many CVOC plumes, biodegradation stalls at cDCE and/or VC. Hence, aerobic degradation of the low chlorinated metabolites (Hanson and Brusseau 1994; Verce et al. 2000, 2001) in an aerobic plume fringe is regarded as an important degradation pathway. Whether or not also anaerobic oxidative degradation in less reducing environments such as the plume front plays a significant role at real field sites is not yet answered and has been addressed in various laboratory experiments (Bradley and Chapelle 1996, 1997; Bradley et al. 1998; Chapelle and Bradley 2003).

Substantial concentrations of the end products of reductive CVOC degradation (ethene and ethane) – i.e. complete degradation of PCE or TCE – may be found exclusively in areas where elevated concentrations of dissolved organic carbon (DOC) are still present and a strongly reducing environment prevails. Down-gradient of these areas, ethene and ethane concentrations usually drop below the detection limit, suggesting that the non-chlorinated end products are rapidly mineralized to CO_2 in a more oxidizing environment.

Given the above mentioned facts that CVOCs are fairly soluble, less degradable than e.g. petroleum hydrocarbons and have a low tendency to adsorb to the soil matrix, causes CVOC contaminations usually to form spatially expanded plumes with a length up to several kilometres.

22.4.2 *Evaluation of Natural Attenuation Potential and Challenges at Chlorinated Hydrocarbon Contaminated Sites*

Basic to application of MNA is a detailed understanding of the localization of free phase (source) and the 3-dimensional spreading of the CVOC plume in the groundwater. Hence, innovative investigation technologies and especially a depth-related groundwater sampling are essential at DNAPL contaminated sites (Held 2007). Regularly (e.g. annual) mapping of the CVOC plume on the basis of analytical results from groundwater monitoring wells is an important tool to describe and monitor spatial changes of the plume extension over time (i.e., to determine whether the plume extension is stable or not). Besides the metabolites cDCE and VC, the end products of reductive dechlorination, i.e., ethene and ethane, also have to be routinely analyzed and mapped.

Concomitant with the investigation of CVOC concentrations, the parameters indicating the biogeochemical environment should be analyzed. These redox indicators (see Section 22.2.3) comprise dissolved oxygen, nitrate, dissolved iron and manganese, sulphate and methane as well as the total and/or dissolved organic carbon (TOC/DOC) content. Apart from conventional analyses of these parameters, redox sensitive textile tapes (RST) may be used which, after incubation, show specific changes in colour due to the redox environment to which they were exposed. The RST may show that the vertical extents of the transition zones are a few centimetres only. Due to the fact that during biodegradation the microorganisms consume the electron acceptors in the sequence given above, different redox zones form within the aquifer. In theory, these redox environments are in the form of interleaved blisters. In reality, also other forms may be observed due to subsurface heterogeneity. Frequently, a complete redox sequence from methanogenic (near the source) to aerobic at the plume fringe can be found. If at all, aerobic conditions can only be found at the plume front. Based on the mapped redox environments and the CVOC distribution, it may be derived in which zones biodegradation (either reductive or oxidative) can occur and in which zones CVOC degradation is not likely.

With regard to possible contributions of the organic material to the supply of H_2 (electron donor) required for the reductive dechlorination of CVOCs, it is helpful to analyze for anthropogenic contaminants as well as for dissolved organic carbon (DOC). Both naturally occurring organic carbon (e.g. peat or organic carbon rich fluidal sediments) or organic carbon from anthropogenic contaminations (e.g. petroleum hydrocarbons) can act as an electron donor to produce H_2. However the microbial degradability of organic carbon depends on the type of DOC. Investigations on the degradability of the DOC have been performed (Rectanus et al. 2005), but these do not consider anaerobic degradation. Furthermore, numerous thermodynamic parameters determine which part of the DOC ends up in methanogenesis and which part may be used for reductive dehalogenation (Ballapragada et al. 1997). Hence, at present the degree of dechlorination cannot be predicted via characterization of the DOC.

Due to the high sensitivity of the redox environments on the degradation processes of CVOCs, factors or measures influencing the redox environments might have significant effects on the CVOC degradation. This might be – among other things – hydraulic alteration of the flow characteristics, either anthropogenically (e.g. by Pump and Treat measures) or naturally (e.g. due to river or sea water level variations) as well as source remediation measures.

Pump & Treat might interrupt the supply of CVOCs into the plume as well as the input of organic non-chlorinated material, altering the redox and degradation regime down-gradient and within the vicinity of the hydraulic measure. If positive or negative effects on the CVOC degradation prevail, depends among other things on the design of the Pump & Treat measure (positioning relative to the source and plume, pumping rate etc.) as well as on the composition of the uncontaminated groundwater (e.g. presence of DOC) up-gradient of the source and laterally next to the pumping well and on the prevalent CVOC contaminants (e.g. high- or low-chlorinated CVOCs).

Also naturally occurring alterations of the groundwater flow characteristics, e.g. due to seasonal or event driven variations of surface water levels, may cause temporal lateral deflection of groundwater flow leading to a meandering of the plume. In case the plume front as well as lateral areas show aerobic conditions, the staggering will lead to a spatially more extensive aerobic zone at the plume fringes. Hence the aerobic bio-reactive area is substantially increased, thus conditions for aerobic CVOC oxidation improve but reductive dechlorination is hampered.

Besides groundwater flow variations, also remediation of other contaminations which serve as electron donors (e.g. BTEX) has substantial impact on the CVOC degradation, as necessary reaction partners are eliminated. Thus, remediation e.g. of an adjacent petroleum hydrocarbon contamination, should be aligned with the remediation of the CVOC contamination.

For evaluation of NA at a chlorinated hydrocarbon contaminated site, the stepwise approach as presented in Section 22.2.2 can be applied, using the methods summarized in Section 22.2.3. Another important proof concerning the sustainability of NA at a CVOC contaminated site relates to the question whether the DOC source is as long lasting as the CVOC source. Sustainability of the biodegradation processes is only provided when the longevity of the electron donor source is at least within the same range as the longevity of the CVOC source. If the source of the electron donors is geogene and not due to anthropogenic spills, it may be assumed that surplus electron donors are available. A more detailed evaluation of this phenomenon in given by Chapelle et al. (2007). In case of an anthropogenic DOC source, the lifetime and source emission for both sources (DOC and CVOC) have to be estimated. Analytical and/or numerical multiphase models are important tools for these predictions.

22.4.3 Enhanced Natural Attenuation

To improve the conditions for reductive CVOC degradation one frequently used ENA measure is to increase the supply of organic matter. Especially degradation of the parental products (TCE, PCE) is stimulated, however also cDCE and VC degradation might be increased by addition of organic material. For this purpose a variety of substrates have been tested or designed. Low cost substrates like molasses or lactate or so-called slow release contaminants like emulsified vegetable oil seem to be promising. In any case the application of these substrates is not trivial. Special technical protocols have been developed providing guidelines to optimize the organic matter supply (Suthersan and Payne 2005; Suthersan et al. 2002). Such injections of organic material may enhance CVOC degradation substantially.

22.5 Natural Attenuation at Tar Oil Contaminated Sites

22.5.1 Introduction

Tar oils are mixtures of hundreds to thousands of different organic contaminants. The main components are polycyclic aromatic hydrocarbons (PAH), NSO-heterocyclic contaminants (NSO-HET), mono aromatic contaminants like benzene, toluene, xylene, and ethylbenzene (BTEX) and short chained alkyl phenols (SCAP). These organic liquids were produced in high quantities as a by-product of town gas production (Collin and Höke 2005). Today coke processing is still a source for tar oil. The composition as well as the properties of tar oil is mainly influenced by the raw material (e.g., coal or lignite) and process parameters (e.g., treatment temperature during production).

During the production of town gas, large amounts of tar were deposed at the sites, causing severe environmental problems due to their toxic and carcinogenic properties. By moving into the subsurface, the tar posed a high risk to soil and groundwater contamination.

During the 19th century the knowledge of organic chemistry increased and over several decades tar oil was a very important raw material for production of dye (coal tar dyes), pharmaceutics, pesticides, resigns, and lacquer. Later it was replaced by crude oil. Creosote, composed of coal tar distillates, is used to treat wood products (preservation) such as railroad sleepers and telegraph poles. It has strong antifungal properties and operates as a impregnate.

Decontamination of such sites is a challenging issue because remediation is time consuming due to various reasons and requires high costs. Tar oil based contaminants were chosen in this Natural Attenuation chapter as a third important class of chemicals. These contaminants are known to be (in contrast to most of the petroleum hydrocarbons and the CVOCs) highly persistent.

22.5.2 Characteristics of Tar Oil

The water solubility of most contaminants is very low, leading to formation of a separate phase when tar oil comes into contact with water. Due to the – in most cases – higher density of tar oil compared to water, the organic phase is able to move downwards through the aquifer and will eventually reach the basis of the aquifer. Such organic liquids having a higher density than water are called dense non-aqueous phase liquids (DNAPL). Depending on the production process the variety of the composition of tar oil is high and some products may even have a lower density than water and therefore form LNAPL (Collin and Höke 2005).

Due to low water solubility and slow dissolution kinetics, NAPLs are long-lasting in groundwater environments. Eberhardt and Grathwohl (2002) calculated, based on laboratory experiments, the time scale for dissolution of tar oils in saturated zones under natural conditions. According to these calculations, tar oil distributed in an aquifer as blobs will release PAHs from decades up to hundreds of years while the dissolution from pools will take hundreds up to thousands of years.

22.5.2.1 Tar Oil Components

Usually, the so-called 16 US EPA PAHs are investigated at tar oil contaminated sites. Additionally, two methylnaphthalenes (1-methylnaphthtalene and 2-methylnaphthalene) are nowadays often included in monitoring programs, because they were found in high quantities in groundwater samples. These contaminants represent a large number of different polycyclic aromatic contaminants with very different physico-chemical and toxic properties. For instance the water solubility varies from 30 mg/L for naphthalene to 3×10^{-3} mg/L for benzo(a)pyrene. For details about the group of PAH see also Table 22.2.

Recently the NSO-heterocyclic (NSO-HET) constituents have been a focus of investigation (Blotevogel et al. 2008, Werner et al. 2008a; Zamfirescu and Grathwohl 2001). The chemical structure of NSO-HET is similar to the PAHs with the exception that one carbon atom of the aromatic ring structure in these organic molecules is replaced by a nitrogen (N), sulfur (S) or oxygen (O) atom. Their

Table 22.2 Selected constituents of tar oil and their physico-chemical properties

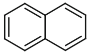
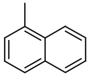
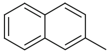
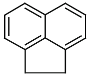
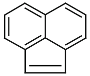
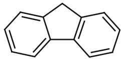
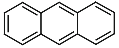
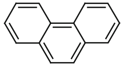
Substance [CAS-No.]	Chemical structure	Melting point [°C]	Boiling point at 1.013 hPa [°C]	Solubility at 25° C [mg/L]	Partitioning coefficient (octanol-water) log k_{ow} [–]	Average content in tar oil[1] [%]	Toxicity[2]
PAH							
Naphthalene [91-20-3]		80	218	31	3.36	0.68	C,(M),(G)
1-Methylnaphthalene [90-12-0]		−30.5	244	25.8	3.87	1.4	(C)
2-Methylnaphthalene [91-57-6]		34.6	243	24.6	3.86	1.77	–
Acenaphthene [83-32-9]		95.3	277.2	3.57	3.92	1.6	E
Acenaphthylene [208-96-8]		93	270	16.1	4.07	2.0	M,G
Fluorene [86-73-7]		115	298	1.98	4.18	1.42	M,G
Anthracene [120-12-7]		218	339.9	0.043	4.45	0.3	(M),(G)
Phenanthrene [85-01-8]		100.5	338.4	1.15	4.5	6.77	M,G

Table 22.2 (continued)

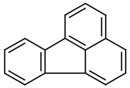
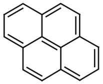
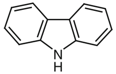
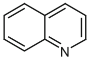
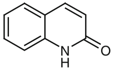
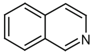
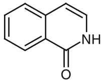
Substance [CAS-No.]	Chemical structure	Melting point [°C]	Boiling point at 1.013 hPa [°C]	Solubility at 25° C [mg/L]	Partitioning coefficient (octanol-water) log k_{ow} [–]	Average content in tar oil[1] [%]	Toxicity[2]
Fluoranthene [206-44-0]		111	383.5	0.26	4.95	4.39	C,M,G,E
Pyrene [129-00-0]		150	393.5	0.135	4.88	2.3	(M),(G),E
N-HET							
Carbazole [86-74-8]		240	350	1.8	3.72	1.5	(C),(M),(G)
Quinoline [91-22-5]		−15	238	6100	2.03	0.3	C,M,G
Quinoline-2-on [59-31-49]		199.5	333	1050[3]	1.26	Metab.	M,G
Isoquinoline [119-65-3]		26	243	5000	2.08	0.2	M,G
Isoquinoline-1-on [491-30-5]		212.5	332.9	480	1.42	Metab.	–

Table 22.2 (continued)

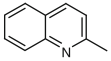
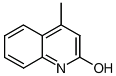
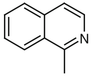
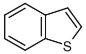
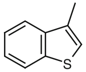
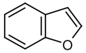
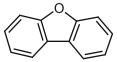
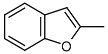
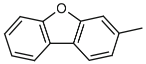
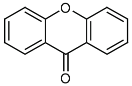
		Melting point	Boiling point at 1.013 hPa	Solubility at 25° C	Partitioning coefficient (octanol-water) log k_{ow}	Average content in tar oil[1]	
Substance [CAS-No.]	Chemical structure	[°C]	[°C]	[mg/L]	[–]	[%]	Toxicity[2]
2-Methylquinoline [91-63-4]		−2	247	499	2.23	≥0.5	C,M,G,E
4-Methylquinolin-2-one [607-66-9]				1760	1.70	Metab.	G
1-Methylisoquinoline [1721-93-3]		10	252	410	2.69	–	M,G
S-HET							
Benzothiophene [95-15-8]		30	222	130	3.12	0.3	(G),E
3-Methylbenzothiophene [1455-18-1]		36	258.1	49	3.54	–	E
O-HET							
Benzofuran [271-89-6]		–18	174	678	2.67	0.1	C,M,G

Table 22.2 (continued)

Substance [CAS-No.]	Chemical structure	Melting point [°C]	Boiling point at 1.013 hPa [°C]	Solubility at 25° C [mg/L]	Partitioning coefficient (octanol-water) log k_{ow} [–]	Average content in tar oil[1] [%]	Toxicity[2]
Dibenzofuran [132-64-9]		80	288	3.1	4.12	1.3	M,G
2-Methylbenzofuran [4265-25-2]		16.3	197.5	160	3.22	–	–
2-Methyldibenzofuran [60826-62-2]		45	305.5	3.2	4.26	–	E
Xanthenon [90-47-1]		174	351	4.52	3.39	Metab.	–

–no data

[1] metab. = metabolite

[2] *C* carcinogen, *G* genotoxic, *M* mutagenic, *E* high ecotoxicity (LC50 <1 mg/L); data in brackets: potentially toxic but not proven

[3] at 20° C

proportion in, for example, creosote is in the range of 3–15%. Some of these contaminants show a relatively high water solubility and therefore they are more mobile than PAHs, hence they may represent up to 40% of creosote-born contaminants in the dissolved fraction. Up to now they are seldom analyzed at contaminated sites, but monitoring campaigns have proven their significance.

The number of contaminants belonging to this group is very high. Therefore it is almost impossible to analyze each single contaminant. Hence, based on their occurrence at contaminated sites (preferentially plumes) and their properties (physico-chemical, biodegradability under in-situ conditions, (eco)toxicity) a list of 20 priority substances was derived for future monitoring (Blotevogel et al. 2008) (see Table 22.3). Still our knowledge of the fate of NSO-HET in the subsurface is scarce.

Table 22.3 Priority list of NSO-HET contaminants

No.	substance [CAS-Nr.]	No.	substance [CAS-Nr.]
	N-HET		O-HET
1	Acridinon [578-95-0]	12	Benzofuran [271-89-6]
2	Carbazol [86-74-8]	13	Dibenzofuran [132-64-9]
3	Chinolin [91-22-5]	14	Methylbenzofurane (2-Methylbenzofuran) [4265-25-2]
4	Chinolinone (Chinolin-2-on) [59-31-49]	15	Methyldibenzofurane (1-, 2-, 4-Methyldibenzofuran) [60826-62-2]
5	Dimethylchinoline (2,4-Dimethylchinolin) [1198-37-4]	16	Dimethylbenzofurane (2,3-Dimethylbenzofuran) [3782-00-1]
6	Isochinolin [119-65-3]	17	Xanthenon [90-47-1]
7	Isochinolinone (Isochinolin-1-on) [491-30-5]		S-HET
8	Methylisochinoline (1-Methylisochinolin) [1721-93-3]	18	Benzothiophen [95-15-8]
9	Methylchinolinone (4-Methylchinolin-2-on) [607-66-9]	19	Methylbenzothiophene (3-Methylbenzothiophen) [1455-18-1]
10	Methylchinoline (2-Methylchinolin) [91-63-4]	20	Hydroxybiphenyl * (2-Hydroxybiphenyl) [90-43-7]
11	Phenanthridinon [1015-89-0]		

Blotevogel et al. (2008), Werner et al. (2008a)
*degradation product of dibenzothiophene

22.5.3 Natural Attenuation Potential of Tar Oil

The most important Natural Attenuation process for organic contaminants is biodegradation. It is known that a lot of different organic contaminants such as BTEX or hydrocarbons, etc., are well degradable even under anaerobic in situ conditions by autochthonous microorganisms. The biodegradability of PAH is also investigated intensively. Most of these studies were performed using single contaminants and/or pure cultures in batch experiments. These results are not transferable to tar oil contaminated sites as the situation there differs considerably. The multicomponent mixture tar oil may have a huge impact on the biological degradability of individual components. In some cases co-metabolic effects may enhance degradation of recalcitrant contaminants. But, it is also possible that biodegradation may be inhibited due to, for example, toxic effects.

Several microorganisms are able to consume 2- to 4-ring PAHs as a carbon source and gain energy from this process. Additionally co-metabolic degradation of some PAHs was observed. Because molecules with lower water solubility are less bioavailable, they are less biodegradable and the degradation rates are low. Therefore degradation of these high molecular weight PAHs is difficult to monitor.

Besides bioavailability, the degradability depends also on the electron acceptors available. Aerobic biodegradation of PAHs is the most effective and the best understood degradation process. The biodegradability of low molecular PAHs using nitrate, ferric iron, manganese(IV), or sulphate is also observed. Nevertheless there is less information available. Only for some contaminants (e.g., naphthalene) the degradation pathways are investigated in detail (e.g., Annweiler et al. 2002; Safinowski and Meckenstock 2006). In general, the degradation rates for PAH are low compared to, for example, mono aromatic contaminants.

The biodegradation of NSO-HET is not yet well investigated. It is known that N-, and O-heterocyclic contaminants better degrade compared to S-heterocyclic contaminants. In some cases an inhibition of aerobic PAH degradation was identified when NSO-HET are available and vice versa (Meyer and Steinhart 2000; Werner et al. 2008a). Similar to PAH the degradability is influenced by the kind of available electron acceptors. The contaminants are likely to be degradable under aerobic conditions. Table 22.4 gives an overview of the degradability under different redox conditions (Werner et al. 2008a).

Substituents and their position in the molecule have a substantial influence on the degradability. If an alkyl group (e.g., methyl group) is located in the direct neighbourhood or is directly bounded to the hetero atom, it will reduce the biodegradation considerably.

Some recalcitrant NSO-HET contaminants are dimethylbenzofuran, methylbenzo(b)furan and benzothiophene (Werner et al. 2008a, b). The length of the plume of these contaminants is comparable or may even be longer than e.g. an acenaphthene plume.

It has been recognized that polycyclic aromatic hydrocarbons may form nonextractable residues (neR) in soil and that this process may be stimulated by

Table 22.4 Degradability of some NSO-HET under different redox conditions

	Electron acceptor				
Substance	O_2	NO_3^-	Fe(III)	SO_4^{2-}	CO_2
Acridinone	?	?	?	?	?
Benzofuran	+	–	+	+	–
Benzothiophene	$+^c$	$+^c$	$+^c$	$+^c$	$+^c$
Carbazole	+	+	+	+	?
Quinoline	+	+	+	+	+
Quinolinone	+	+	+	+	+
Dibenzofuran	+	+	+/–	+	?
Dimethylbenzofuran	+	–	–	?	?
Dimethylquinoline	+	–	–	–	?
2-Hydroxybiphenyl	+	?	?	+	+
Isoquinoline	+	+	–	+	+
Isoquinolinone	+	–	+	+	+
Methylbenzofuran	+	+	–	?	?
Methylbenzothiophene	+	?	?	?	?
Methylquinoline	+	+/—	+/—	+/—	+/—
Methylquinolinone	+	—	+/—	+/—	?
Methyldibenzofuran	?	?	?	?	?
Methylisoquinoline	?	?	?	?	?
Phenanthridinon	?	?	?	?	?
Xanthenon	?	?	?	?	?

+ transformation proven; – transformation not proven; +/— transformation not always or not for all isomers proven; ? no data available; *c* transformation only co-metabolic proven

microbial activities (Eschenbach et al. 1998). In this case the PAHs are not mineralized but transformed to polar metabolites. These products may form stable bonding to natural organic matter such as humic substances. This so-called humification was first investigated for aerobic PAH degradation and was recently also proven for anaerobic conditions using ^{14}C-labled PAH and NSO-HET contaminants (Berghoff et al. 2007).

The mobility of organic contaminants in the subsurface is influenced mainly by sorption processes on natural organic matter. PAHs are hydrophobic substances and, therefore, adsorb much stronger than e.g. monoaromatic contaminants. This is indicated by the high octanol-water partition coefficient k_{ow} (see Table 22.2). As a consequence, the high molecular PAHs (e.g. Benzo[a]pyrene) are found only in the source zone and usually not in the plume. Due to the fact that sorption is reversible later mobilisation might be possible.

22.5.4 Summary

Tar oils are multi-component mixtures. Without total source removal the emission of contaminants out of the NAPL phase will go on for a long time. The mobility as well as the degradability of the constituents differs in a wide range (e.g. BTEX vs. high molecular PAH) and is on the one hand substance specific and on the other hand it might be influenced by the multi-component mixture itself. The biodegradation rate under in-situ conditions is very low for a lot of tar oil contaminants. For some substances the degradability especially under anaerobic conditions is even not known.

Nevertheless, for most of the sites the plume length does not exceed a few hundred meters even if the sites were closed decades ago. This fact indicates the potential of the autochthonous microorganisms for degrading these contaminants and/or for effective sorption processes taking place. Additionally, knowledge about the in-situ processes increased considerably. For identification and quantification of the Natural Attenuation processes the strategy and the methods described in Sections 22.2.2 and 22.2.3 are recommended.

In several case studies the potential of Natural Attenuation processes was investigated in detail and implemented (e.g. Neuhauser et al. 2009; Rügner et al. 2004; Werner et al. 2008a). For the acceptance of Monitored Natural Attenuation at tar oil contaminated sites, an intensive discussion with the local authorities is necessary. Long lifetimes of tar oil sources and the uncertainties in the degradation processes for PAHs and NSO-HET are the outstanding challenges.

22.6 Conclusions and Outlook

In recent years, substantial work has been undertaken in the field of research and development, preparation of technical guidelines, and legal implementation of MNA, mainly in the USA, but also in various European nations. This work and especially the comprehensive six year priority research program performed in Germany until the end of 2008 and involving more than 1000 researchers, have contributed to the nowadays highly improved understanding of physical, chemical, and biological processes that contribute to the transport and degradation of contaminants in soil and groundwater. In this context, it was important to realize that some of the most important processes are restricted to only small spaces within the huge mass of contaminated soil and groundwater. Usually the inner parts of contaminant plumes are depleted of mobile electron acceptors and therefore the biodegradation predominantly takes place in the fringe of the plume body, apart from degradation processes using immobile electron acceptors as Fe(III)/Mn(IV) or methanogeneses, respectively.

Besides an increase of process understanding, the progress was also accompanied by intensified development of methods, needed to investigate NA processes and especially to derive quantitative and not only qualitative data from site samples. Whereas in the past it was quite easy to determine biodegradation rates in laboratory experiments, today this is for many contaminants also possible in the field, e.g., applying the contaminant specific isotope analysis (CSIA). The method can be used to determine degradation rates for productive degradable carbon contaminants (e.g., (mono)aromatics) as well as for the degradation of chlorinated hydrocarbons which are usually used as electron acceptors. Furthermore, DNA-based methods are emerging more and more. At present they are used, for example, for the detection of *Dehalococcoides ethenogenes*, the only microorganism known to be able to completely mineralize CVOCs.

Despite the huge efforts spent in understanding the processes, the existing individual national protocols are not yet harmonized (Rügner et al. 2006). The most challenging questions are whether MNA can be used as a stand-alone approach, whether source removal should be mandatory and whether MNA is effective enough to reduce contaminant concentrations below the target levels within a manageable time frame, which is usually 30 years. This understanding ignores many of the inevitable processes. Most sources have a lifetime larger than human lifetimes, even for such comparable moderate soluble contaminants like (mono)aromatics, it might take hundreds of years until sources comprising small NAPL-blobs are dissolved (Eberhardt and Grathwohl 2002). In case the source is removed by the common applied technique, Dig and Dump, the soil exchange must be complete. In case of incomplete removal of NAPLs, it needs only 10 cm of residual phase being passed by the natural groundwater flow to increase the contaminant concentration by dissolving again to values comparable to its maximum solubility.

Although research on NA started with overwhelming enthusiasm, maybe even longing for remediation activities becoming dispensable, today the scientific and regulatory communities have a realistic understanding about the possibilities and limitations of NA. This results in the fact that MNA (provided it is effective at the site of concern) is considered as one possible approach in the area of Brownfield management. Only in a small number of cases (e.g., where the source cannot be removed or treated because the respective area is overbuilt or not accessible due to other reasons) MNA is accepted as a stand alone approach. In most cases it is used as an additional step after active remediation techniques have failed to reach the permissible contaminant levels in the predicted time frame and prolongation is uneconomical and disproportional. MNA as a stand alone approach might be regarded as an option for e.g. mega sites which are not remediable within an appropriate temporal and financial frame applying conventional remediation technologies.

Using MNA in the described context seems to be quite reasonable, because it appreciates on the one hand that groundwater is an important source of drinking water in many regions of the world and should be managed in a sustainable way, to guarantee water resources for future generations. On the other hand, this approach also considers natural and technical constraints such as the impossibility at many sites of cleaning up to the permissible contaminant levels with active remediation techniques.

Considering that in many cases Dig and Dump procedures do not remove the contamination completely, the discussion came up as to what extent source removal is necessary and MNA is effective enough to do the rest of the job (contaminant removal). This must still be concordant with the results of a Risk Assessment: the contamination must be diminished to an extent that after source removal, sensitive receptors are no longer affected. This approach, however, is in some conflict with the groundwater precaution policy. Hence, in the field, where MNA is applicable or not there is still some discussion needed.

Because the application of MNA is a knowledge intensive process it seems necessary to increase efforts in further training of practitioners and regulators not dealing daily with MNA. To enhance the acceptance of MNA, it would also help to supply a database (e.g., on the internet) with successful MNA applications.

References

AFCEE (2000) Designing monitoring programs to effectively evaluate the performance of natural attenuation. US Airforce Center for Engineering and the Environment

American Society for Testing and Materials (2004) STM E1943 – 98(2004) Standard guide for remediation of ground water by natural attenuation at petroleum release sites

Annweiler E, Michaelis W, Meckenstock RU (2002) Identical ring cleavage products during anaerobic degradation of naphthalene, 2-methylnaphthalene, and tetralin indicate a new metabolic pathway. Appl Environ Microbiol 68:852–858

Ballapragada BS, Stensel HD, Puhakka JA, Ferguson JF (1997) Effect of hydrogen on reductive dechlorination of chlorinated ethenes. Environ Sci Technol 31:1728–1734

Berghoff A, Wienberg R, Farahbakhsh M, Dethlefsen F, Mahro B (2007) Comparison and evaluation of site-specific microcosm-experiments to identify and kinetically describe the processes that determine the microbial degradation of PAH in groundwater (in German). Final Report, Priority Program KORA, Project 2.7, FZK 02WN0436

Blotevogel J, Reineke AK, Hollender J, Held T (2008) Identification of NSO-heterocyclic priority substances for investigating and monitoring creosote-contaminated sites (in German). Grundwasser 13:147–157

Bockelmann A, Ptak T, Teutsch G (2000) Field scale quantification of contaminant mass fluxes and natural attenuation rates using an integral investigation approach. In: Bjerg PL, Engesgaard P, Krom TD (eds) Groundwater 2000. Balkema, Copenhagen, Denmark, pp 309–310

Bradley PM (2003) History and ecology of chloroethene biodegradation: a review. Bioremed J 7:81–109

Bradley PM, Chapelle FH (1996) Anaerobic mineralization of vinyl chloride in Fe(III)-reducing aquifer sediments. Environ Sci Technol 30:2084–2092

Bradley PM, Chapelle FH (1997) Kinetics of dichloroethylene and VC mineralization under methanogenic and Fe(III)-reducing conditions. J Contamin Hydrol 31:111–127

Bradley PM, Landmeyer JE, Dinicola RS (1998) Anaerobic oxidation of [1,2-14C]dichlorethene under Mn(IV)-reducing conditions. Appl Environ Microbiol 64:1560–1562

Chapelle FH, Bradley PM (2003) Redox conditions and the reductive/oxidative biodegradation of chlorinated ethenes in groundwater systems. Dehalogenation: Microbial Processes and Environmental Applications, Kluwer Academic, Norwell, pp 373–384

Chapelle FH, Novak J, Parker J, Campbell BG, Widdowson MA (2007) A framework for assessing the sustainability of monitored natural attenuation. U.S. Geological Survey, Circular 1303

Collin G, Höke H (2005) Tar and Pitch. In: Ullmann's encyclopedia of industrial chemistry, Wiley-VCH, Weinheim

Davis JW, Carpenter CL (1990) Aerobic biodegradation of vinyl chloride in groundwater samples. Appl Environ Microbiol 56(12):3878–3880

Domenico PA (1987) An analytical model for multidimensional transport of a decaying contaminant species. J Hydrol 91:49–58

Eberhardt C, Grathwohl P (2002) Time scales of organic contaminant dissolution from complex source zones: coal tar pools vs. blobs. J Contamin Hydrol 59:45–66

Eschenbach A, Wienberg R, Mahro B (1998) Fate and stability of nonextractable residues of (14C)PAH in contaminated soils under environmental stress conditions. Environ Sci Technol 32:2585–2590

Grandel S, Dahmke A (2008) Guideline– natural attenuation at chlorinated hydrocarbon contaminated sites. Methods, recommendations and comments for NA investigation and evaluation (in German). Institut für Geowissenschaften, Universität Kiel, Kiel

Grathwohl P (1998) Diffusion in natural porous media: contaminant transport, sorption, desorption and dissolution kinetics. Kluwer Academic, Boston, 207 pp

Hanson RS, Brusseau GA (1994) Biodegradation of low-molecular-weight halogenated organic compounds by aerobic bacteria. In: Chaudhry GR (ed) Biological degradation and bioremediation of toxic chemicals. Chapman and Hall, London, pp 277–297

Hata J, Takamizawa K, Miyata N, Iwahori K (2003) Biodegradation of cis-1,2-dichloroethene and vinyl chloride in anaerobic cultures enriched from landfill leachate sediment under Fe(III)-reducing conditions. Biodegradation 14:275–284

Held T (ed) (2007) Controlled natural attenuation of contaminants with regard to remediation of contaminated groundwater and soil. Microbiological NA procedures: technical basis for the use of methods for the assessment of the natural microbial degradation of contaminants in an aquifer (in German). ProcessNet/KORA (März 2007)

Hunkeler D, Andersen N, Aravena R, Bernasconi SM, Butler BJ (2001) Hydrogen and carbon isotope fractionation during aerobic biodegradation of benzene. Environ Sci Technol 35: 3462–3467

King MWG, Barker JF, Devlin JT, Butler BJ (1999) Migration and natural fate of a coal tar creosote plume 2. Mass balance and biodegradation indicators. J Contamin Hydrol 39:281–307

LABO (2009) Considering natural attenuation processes for contaminated site management (in German). Bund/Länder Arbeitsgemeinschaft Bodenschutz (German federal and state panel on soil protection)

Laitine J (2006) In-situ soil and groundwater bioremediation techniques and applications, PhD thesis, Tampere Polytechnic, University of Applied Sciences (TAMK), Finnland

Meyer S, Steinhart H (2000) Effects of heterocyclic PAHs (N, S, O) on the biodegradation of typical tar oil PAHs in a soil/compost mixture. Chemosphere 40:359–367

Michels J, Stuhrmann M, Frey C, Koschitzky H-P (2008) Guidelines – natural attenuation for contaminated site management. Evaluation and application, legal and economical aspects, and public acceptance (in German), Dechema e.V., Frankfurt

Miles B, Peter A, Teutsch G (2008) Multicomponent simulations of contrasting redox environments at an LNAPL contaminated site. Ground Water 46(5):727–742

Neuhauser EF, Ripp JA, Azzolina NA, Madsen EL, Mauro DM, Taylor T (2009) Monitored natural attenuation of manufactured gas plant tar mono- and polycyclic aromatic hydrocarbons in ground water: a 14-year field study. Ground Water Monitor Remed 29(3): 66–76

NRC (2000) Natural attenuation for groundwater remediation, National Research Council, National Academy, Washington, USA

Pennington JC, Gunnison D, Harrelson DW, Brannon JM, Zakikhani M, Jenkins TF, Clarke JU, Hayes CA, Myers T, Perkins E, Ringelberg D, Townsend D, Fredrickson H, May JH (1999) Natural attenuation of explosives in soil and water systems at Department of Defense sites: Interim report. Technical report EL-99-8, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS

Peter A, Miles B, Teutsch G (2008) Estimation of emission from an LNAPL contaminated zone considering groundwater recharge. Environ Geol 55(2):321–337

Peter A, Miles B, Vigelahn L, Teutsch G (2006) Evaluation of natural attenuation at the petroleum hydrocarbon contaminated site Niedergörsdorf, Tanklager 1 (in German). Technical Information of the Environmental Protection Agency of the State Brandenburg, vol 12. http://www.mugv.brandenburg.de/cms/media.php/lbm1.a.2334.de/labo_nr12.pdf. Accessed 18 August 2010

Potter TL, Simmons KE (1998) Composition of petroleum mixtures, vol 2. Total Petroleum Hydrocarbon Criteria Working Group

Rectanus HV, Widdowson MA, Novak JT, Chapelle FH (2005) Method development for quantifying potential available organic carbon in sediments. In: Third European bioremediation conference, Chrania, Krete, Greece, 4–7 June 2005, Proceedings, P004

Rice DW, Dooher BP, Cullen SJ, Everett LG, Kastenberg WE, Grose RD, Marino MA (1995) Recommendations to improve the cleanup process for California's leaking underground fuel tanks (LUFTs). UCRL-AR-121762, Lawrence Livermore National Laboratory, Environmental Protection Department

Rügner H, Finkel M, Kaschl A, Bittens B (2006) Application of monitored natural attenuation in contaminated land management – a review and recommended approach for Europe. Environ Sci Policy 9:568–576

Rügner H, Holder T, Ronecker U, Schiffler G, Grathwohl P, Teutsch G (2004) Natural attenuation investigations at the former gas manufacturing plant "Kehl" (in German). Grundwasser 1:43–53

Safinowski M, Meckenstock RU (2006) Methylation is the initial reaction in anaerobic naphthalene degradation by a sulphate-reducing enrichment culture. Environ Microbiol 8:347–352

Stupp HD (2007) Spreading of MTBE and chlorinated hydrocarbons in groundwater: comparison of groundwater transport and plume dimensions. In: Barcello D (ed) The handbook of environmental chemistry. Springer, Berlin Heidelberg

Suthersan SS, Lutes CC, Palmer PL, Lenzo F, Payne FC, Liles DS, Burdick J (2002) Technical protocol for using soluble carbohydrates to enhance reductive dechlorination of chlorinated aliphatic hydrocarbons. F41624-99-C-8032. (eds) Air Force Centre for Environmental Excellence (AFCEE), Brooks, TX, USA; Environmental Security Technology Certification Program (ESTCP), Arlington, VA, USA

Suthersan SS, Payne FC (2005) In situ remediation engineering. CRC, Boca Raton, FL

Teutsch G, Grathwohl P, Schiedek T (1997) Literature review on natural attenuation of contaminants in groundwater (in German), Reports on contaminated site management, vol 35/97, Environmental Protection Department of State Baden-Württemberg, Karlsruhe, Germany

Teutsch G, Rügner H (1999) Quantification of natural attenuation in the field – possibilities and limitations. In: Kreysa G, Track T, Michels J, Wiesner J (eds) Natural attenuation – possibilities and limitations natural of remediation strategies (in German). Dechema, Gesellschaft für Chemische Technik und Biotechnologie e.V., Frankfurt/M, Germany, pp 43–58

US EPA (1994) How to evaluate alternative cleanup technologies for underground storage tank sites: a guide for corrective action plan reviewers. US Environmental Protection Agency, Washington, DC

US EPA (1998) Technical Protocol for evaluating natural attenuation of clorinated solvents in ground water. EPA/600/R-98/128

US EPA (1999) Use of monitored natural attenuation at superfund, RCRA corrective action, and underground storage tank sites. 9200.4-17P, U.S. EPA – Office of Solid Waste and Emergency Response, Washington, DC

US EPA (2005) A decision-making framework for cleanup of sites impacted with light non-aqueous phase liquids (LNAPL), EPA Office of Research and Development, Washington, DC

US EPA (2009) Commonly asked questions regarding the use of natural attenuation for petroleum contaminated sites at federal facilities. http://www.epa.gov/fedfac/documents/petrol.htm. Accessed 18 August 2010

Valo R, Apajalahti J, Salkinoja-Salonen M (1985) Studies on the physiology of microbial degradation of pentachlorophenol. Appl Microbiol 21(5):313–319

Verce MF, Ulrich RL, Freedman DL (2000) Characterization of an isolate that uses vinyl chloride as a growth substrate under aerobic conditions. Appl Environ Microbiol 66: 3535–3542

Verce MF, Ulrich RL, Freedman DL (2001) Transition from cometabolic growth-linked biodegradation of vinyl chloride by a Pseudomonas sp. isolated on ethene. Environ Sci Technol 35:4242–4251

Wabbels D, Teutsch G (2008) Guideline – natural attenuation at mineral oil contaminated sites. Procedures, recommendations and comments with regard to research and appraisal (in German). Universität Tübingen, Tübingen

Werner P, Börke P, Hüsers N (2008a) Guideline – natural attenuation at tar oil contaminated sites (in German). Schriftenreihe des Institutes für Abfallwirtschaft und Altlasten, TU Dresden, Band 58 ISBN 978-3-934253-50-6

Werner P, Hüsers N, Schönekerl S, Lorbeer H, Börke P, Leibenath C (2008b) Implementation of natural attenuation on a wood preservation facility contaminated with tar oil, Con Soil 2008, LeS: E.6 Natural Attenuation

Wiedemeier TH, Chapelle FH (1998) Technical guidelines for evaluating monitored natural attenuation of petroleum hydrocarbons and chlorinated solvents in groundwater at naval and marine corps facilities, naval facilities. Engineering Command, Alternative Restoration Technology Team

Wiedemeier TH, Rifai HS, Newell CJ, Wilson JT (1999) Natural attenuation of fuels and chlorinated solvents in the subsurface. Wiley, New York

Wiedemeier TH, Wilson JT, Kampbell DH, Miller RN, Hansen JE (1995) Technical protocol for implementing intrinsic remediation with long-term monitoring for natural attenuation of fuel contamination dissolved in groundwater, US Air Force Center for Environmental Excellence, San Antonio, TX

Wilson JT, Wilson BH (1985) Biotransformation of trichloroethylene in soil. Appl Environ Microbiol 49(1):242–243

Zamfirescu D, Grathwohl P (2001) Occurrence and attenuation of specific organic compounds in the groundwater plume at a former gasworks site. J Contamin Hydrol 53:407–427

Part VII
Frameworks

Chapter 23
Bringing Sustainable Management of Contaminated Sites into Practice – The Role of Policy and Regulations

Joop J. Vegter and Harald Kasamas

Abstract This chapter focuses on application of scientific knowledge in dealing with contaminated sites within the broader context of policies and regulations. It reviews the development of strategies in industrial countries and describes the current understanding of how to tackle the problem of contaminated sites in a sustainable manner. Since the first discovery of contaminated sites at the end of the 1970s, public and political perception has changed and the understanding of the nature of the problem has increased considerably. Consequently, strategies for managing these problems have been further developed and improved. Three generations of contaminated sites policy are identified and described in this chapter; from early command-and-control regulations at a national level towards more flexible, site-specific and incentive-driven management approaches at the local level. Today, contaminated sites policy needs to address environmental and spatial planning aspects. It is important to explore and promote solutions in a multi-stakeholder environment that satisfy both environmental and social-economic needs of the society. Internationally accepted concepts that could lead to better problem solutions have been developed jointly in multinational partnerships, like the Risk-Based Land Management (RBLM) Concept of the European Union (EU) network CLARINET. Furthermore, the chapter briefly introduces the general soil protection policy currently under development in the EU. Contamination is one of the identified soil threats in the EU Thematic Strategy for Soil Protection and prevention of new soil contamination should be the key aim for the future in order to provide an added value to already existing national regulations.

J.J. Vegter (✉)
Vegter Advice, Amstelveen, The Netherlands
e-mail: joopvegter@mac.com

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_23,

Contents

23.1 Introduction

This chapter focuses on applications of the scientific knowledge related to assessing and managing contaminated sites, in the broader context of policies and regulations. Dealing with contaminated sites is a cross-cutting issue, where human health and environmental issues and spatial planning and sometimes water management related considerations meet. Assessment of contaminated site problems has often to be performed in complex social environments and solutions must be found that satisfy environmental and social-economic needs. For policy makers, the designers of policy instruments such as regulations, subsidies, tax incentives and special programs for public-private partnership, it is nowadays important to explore and promote solutions to contaminated site problems in a multi-stakeholder environment. Also from the viewpoint of the public authorities, the "regulators", who have to implement policies and make decisions concerning a site-specific solution for a contaminated site problem, a contaminated site is not merely a "technical problem" that can be solved by engineering alone. Perception of the problem by different stakeholders may concern human health or ecological issues, but also financial and social aspects. This also holds true for the solution of a contaminated site problem. Will health risks be sufficiently reduced? Who will bear the responsibilities and liabilities if the solution does not achieve what was planned or when new scientific insights reveal shortcomings of the solution? In view of these complex questions, public authorities should be very focussed on choosing the right Risk Assessment and decision support tools and be aware of their strength and weaknesses and the uncertainties that may remain unaddressed. In risk-based (fitness for use) solutions to contaminated land problems, uncertainties in social and economic domains may even be larger than the uncertainties related to adverse effects of residual contamination, and this may lead to a (temporary) lower priority for improving methods for classical toxicological Risk Assessment.

The complexity of contaminated site problems and their solutions was not perceived immediately after the first discoveries of major incidents of the 1980s, like Love Canal, New York State; Times Beach, Missouri; and Lekkerkerk, the Netherlands. Policy perspectives concerning soil contamination have changed gradually during the almost 30 years that most industrialized countries have addressed these problems. This chapter will review these developments and describe the current generally accepted framework for contaminated site Risk Assessment and management from a policy perspective. It will also briefly introduce the general soil protection strategy that is currently under development in Europe.

Soil contamination is one of the many threats to soil resources that have to be addressed in soil protection, and "historical" or "legacy" contaminated sites are only one part of the general contamination problem, which includes also ongoing and new contamination. However, policies for contaminated sites have already been in place for decades in many industrialised countries. International networks have contributed significantly to the general consensus within the participating European countries on how to deal with contaminated sites. This policy framework is known as Risk-Based Land Management (RBLM) (Vegter et al. 2002). RBLM is firmly rooted in contaminated site Risk Assessment practices from Europe, the US and Canada, but it has the wider perspective of sustainable development, in particular the need to consider the timing of any intervention and the future consequences of any particular solution in relation to environmental, economic, social and cultural dimensions.

23.2 The Development of an Environmental Policy for Soil in the European Union

Compared to other areas of environmental policy, soil protection policy developed much later in most countries than policies for water and air. Some countries do not even have any soil protection policy in place. However, urgency to consider the protection of soil resources has increased. The Millennium Ecosystem assessment (2005) stated that over the past 50 years, humans have changed ecosystems more rapidly and extensively than in any comparable period of time in human history. The services these ecosystems provide are likely to decline with direct and indirect effects on human health and welfare. Soil does play a central role in ecosystem services (including climate regulation for instance) and soil formation is an essential support function for many of them.

Many threats to soil have now been identified. In the European Union (EU) Soil Communication (2002) eight threats to soil were considered important enough to undertake action: erosion, organic matter decline, Biodiversity loss, contamination, sealing, compaction, salinization and floods and landslides. The Soil Communication was followed by a consultation process that included an Advisory Board and five so-called technical working groups with experts from the EU Member States and stakeholder organisations. The concluding reports aimed at a state-of-the-art assessment of the socio-economic drivers leading to environmental pressures on soil, which can result in changes in the state of the soil and to impacts on human health, soil functions and other environmental resources. The reports also explored options for policy responses. This general framework for analysing environmental issues is known as “DPSIR” (Drivers, Pressures, State, Impact and Responses) (Van Camp et al. 2004). The EU consultation process led to a new EU Soil Communication (2006), a proposal for a Soil Framework Directive and an extended impact assessment of various policy options. At present, soil is high on the political agenda of the EU.

23.2.1 The Status of Soil and Soil Contamination

Contamination refers to the present or past introduction (either direct through leakages, spills and applications of contaminated products or indirect through atmospheric deposition or contaminated surface waters) of chemical substances into the soil system by human activities that may lead to adverse effects on land use and soil functions and/or to adverse effects in other parts of the environment.

The invisibility of soil contamination, the heterogeneity of soils, and the fact that sites are often privately owned, make systematic quantitative assessments of the status of contamination in soils costly and hard to enforce. In Europe, reliable spatially representative information is lacking, even in countries where a national contaminated site remediation program is in place, At present only rough estimates about the number of sites that may need remediation exist. In view of the differences in national approaches, even these rough estimates cannot be compared internationally.

The working group on Contamination (Van Camp et al. 2004), one of the technical working groups established under the EU Thematic Strategy for Soil Protection, concluded that a qualitative description of the status of soil contamination in Europe is possible and sufficient to serve as a starting point for policy making at the EU level. The working group considered the following broad categories of land use:

1. Agricultural areas including areas with intensive forestry.
2. Natural areas including recreational areas and areas with extensive forestry.
3. Urban areas and infrastructures.
4. Soils under surface waters or sediments.

In Europe, the overall soil conditions with respect to contamination for each of these land use categories are described below.

23.2.1.1 Agricultural Areas

Soils in agricultural areas are under pressure from atmospheric deposition (acidifying substances, persistent organic pollutants (POPs), heavy metals), direct input of pesticides, manure and other biowaste. These pressures may lead to the slow accumulation of heavy metals and POPs in top soil and to accumulation of phosphate in areas with intensive animal farming.

The use of land for agriculture is endangered in some problem areas, because of safety of food crops. But in general, most agricultural soil is still fit for its use and for future land use changes for what concerns contamination, although a transition from eutrophic farmland to oligotrophic nature may take some time. It should be realised that also a natural succession towards mature terrestrial ecosystems will take its time.

23.2.1.2 Natural Areas

Soils in natural areas share the same pressures from atmospheric deposition as agricultural soils. An additional concern in these areas is related to nitrogen deposition, which may lead to eutrophication and acidification. Soils in natural areas may slowly accumulate POPs and heavy metals but these soils are in general still in a reasonably good shape. Hence the quality of soil would not be a barrier for a change in land use.

23.2.1.3 Urban Areas and Infrastructures

Urban soils and soils related to infrastructure (roads, railways, powerlines, also outside urban areas) are often contaminated. The most important soil contaminants are polyaromatic hydrocarbons (PAHs), lead, copper and zinc from power lines and masts, cadmium from transport, herbicides in the vicinity of roads and railways, asbestos from demolished buildings, mineral oil and volatile organic compounds (VOCs). These soils are often unfit for sensitive land uses like playgrounds for children and vegetable gardens. Another risk category for human health is accumulation of volatile compounds (mostly from severe groundwater contamination) under or in houses. Contaminated soil may lead to contaminated indoor dust and thus to an increased human exposure. Apart from heavily contaminated sites, and problems with volatile compounds under houses due to groundwater contamination, the impacts on human health are generally limited if the soil is not used for vegetable gardens or as playground for children.

23.2.1.4 Sediments

Sediments are the major sinks for water contamination. They reflect the former large-scale contamination of surface waters (metals, mineral oil, PAHs, polychlorinated biphenyls or PCBs, old pesticides). The quality of the surface waters has much improved due to more stringent emission controls, but now the sediments have become a threat for their ecological impacts in the cleaner waters. In many cases, sediments are generally unfit for use on land, after dredging, in agricultural and natural areas. Because river- and harbour- management often requires dredging, the contaminated sediments are a big burden for society. In addition, contaminated sediments have impacts on terrestrial soils after flooding.

The general picture that emerges from the characterisations above is that soils in agricultural and nature areas in Europe are usually in an acceptable state with respect to contamination, but are under pressure. This general picture may be valid also for other parts of the world although it should be realised that diffuse contamination can sometimes affect quite large areas. Taking into account the average land use distribution in 15 EU countries (Fig. 23.1; data elaborated from Eurostat, LUCAS survey (2003)), it may be generally concluded that approximately 70% of EU soils are still in reasonable shape, with the exclusion of problems occurring in certain areas. However, if pressures continue at the current level, as it is already evident in some problem areas, impacts will start to occur on a larger scale, because contaminants quite often tend to accumulate in soils, leading to a degradation of soil

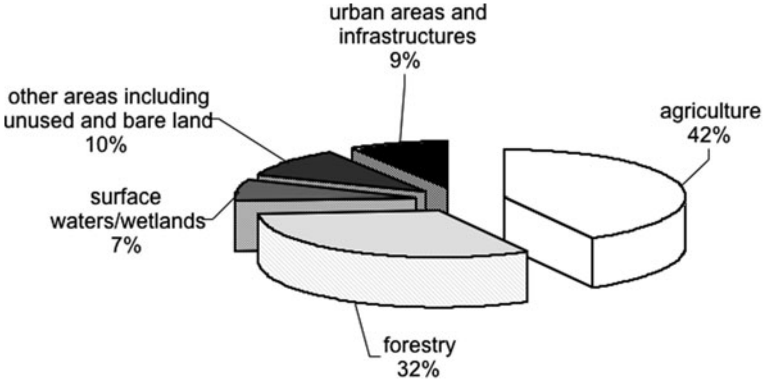


Fig. 23.1 Average land use distribution in 15 EU countries. Data elaborated from Eurostat, LUCAS survey (2003)

quality. So, these pressures should be addressed over longer time periods. On the other hand, many urban soils and sediments are already heavily affected. Prevention should stop further deterioration and the risks of the currently contaminated sites should be adequately managed.

23.2.2 Prevention of Contamination and Management of Contaminated Sites

Although detailed information about the state of soils with respect to contamination is lacking there is enough sound scientific information about the socio-economic drivers, the pressures on land use by human activities and the impacts of a bad soil status to justify the development of a soil protection policy. A few distinctive features that set soil apart from other areas of environmental policy are relevant for the development of soil protection:

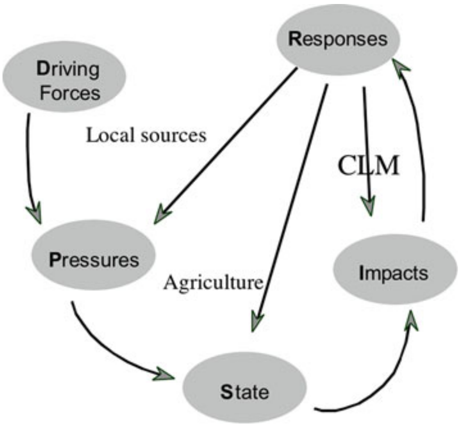
- Soil is a non-renewable resource with potentially rapid degradation rates and extremely slow formation and regeneration processes. Therefore, prevention and precaution should be the core of soil protection policies.
- Since contaminants can exceed irreversibility thresholds unnoticed and become “chemical time bombs” (Salomons and Stigliani 1995), it is essential to have anticipatory policies based on monitoring and early warning systems to protect the environment and human health.
- Since soil is generally submitted to property rights, soil protection policy may make use of the environmental liability of landowners.
- Maintaining soils in good condition is an essential precondition for the long-term sustainability of our society.
- The complexity of local soil contamination is such that avoidance of new contamination has to be “key aim for the future”.
- Soil protection policies should leave enough space for decisions at the local level, because of the geographically diverse nature of soils and the diversity of soil functions and soil uses.

In view of the considerations mentioned above the technical working group on Contamination proposed four specific policy strategies concerning respectively:

- Local sources of contamination.
- Agricultural soil uses.
- Management of contaminated sites.
- Large-scale diffuse contamination.

The strategies are related to the way the land is used and identify the owner/user of the site as the primarily responsible party for the prevention of contamination and the protection of the soil. The strategies can be linked to the DPSIR scheme (Fig. 23.2).

Fig. 23.2 Soil policy strategies (responses) for contamination. Policies for local sources aim at avoiding pressures. For diffuse inputs like agriculture the state of soil should be kept in balance. Contaminated land management (CLM) aims at restoring or reducing impacts



Contamination from local sources, as defined by the working group, corresponds to activities where it is not necessary to put (contaminating) substances into the soil. Oil tanks are not supposed to leak, properly designed waste dumps do not need to have substantial emissions to groundwater and leaching from construction materials can be reduced without affecting their functionality. Prevention should aim to make it virtually impossible for emissions to occur, by containment techniques or precautionary measures at industrial installations. This means that there is no "acceptable emission of contaminants to soil" in these cases, only a risk of failure of preventive devices, which should be "As Low As Reasonably Achievable". Soil is only used for geotechnical support or to provide space for these activities.

On the other hand, activities like agriculture do address the soil as an ecosystem. Agriculture modifies the ecological cycles of energy and matter to adjust them to agricultural use. Sustainable agriculture is not possible without some input of fertilisers. However, to keep the soil in good shape, accumulation of contaminants must be avoided and the inputs must be balanced to the soil system with its normal outputs without adversely affecting other parts of the environment.

For contaminated sites resulting from past industrial activities it is often impossible to apply the "polluters pay principle". In cases that the polluter cannot be legally addressed, the owner of the site should be made responsible for managing and improving the situation, with adequately allocated financial support from public funds. Management and remediation of contaminated sites can be based on Risk Assessment and Risk Management approaches described in the reports of the CARACAS (Ferguson and Kasamas 1999; Ferguson et al. 1998) and CLARINET (Vegter et al. 2002) concerted actions.

A preventive policy approach for the still ongoing large scale diffuse contamination by atmospheric deposition and, in the case of sediments, water contamination, requires large scale integration of soil protection, air and water policies and land use policies. The European Water Framework Directive (2000) provides opportunities for management of water quality and quantity at the river basin scale and may become an important vehicle for soil and sediment protection and further integration of environmental management. The abatement of large-scale diffuse contamination problems will for long be the task of public authorities and international coordination is necessary due to the trans-boundary nature of the environmental problem and its economic repercussions. There is a need for a more proactive planning for the future uses of soil and water resources. Sound management of soil and water resources calls for harmonisation of spatial planning with the environment, so that land uses can be optimised with respect to soil quality and the hydrological and ecological situation.

In 1972 the Council of Europe (European Soil Charter 1972) already stressed the importance of spatial harmonization of land uses and soil functions. More recently, the UN Millennium Ecosystem Assessment (2005) emphasized this again by stating that a spatially diverse "Adapting Mosaic" socio-economic development scenario is best for improvement of Ecosystem Services. In this scenario, harmonisation of socio-economic development with the environment should take place at the regional watershed-scale. This in turn requires that local institutions and local-level political decision-making are strengthened. Local management strategies involving Ecological Services should become more common. The trend to move away from detailed national command and control policies towards more local decision-making based on local soil or land- management plans is indeed observed in countries with a longer tradition in contaminated sites policy making.

23.3 Three Generations of National Contaminated Sites Management Policies

Within the broader framework of soil protection policies, contaminated sites management has a special position. It addresses contamination due to past polluting activities, which is still present in soil and groundwater. There is no equivalent in policies for other environmental compartments like water and air, in which contamination disappears due to dilution within a relatively short timeframe. Therefore, different from the soil, no historical contamination is found in these compartments. Without special remedial actions soil contamination is very persistent, due to the fact that soil is a non-mixing compartment. Current policies in the more experienced countries (e.g., in the EU and in the USA and Canada) stimulate solutions for contaminated site problems that are linked to other socio-economic activities (or problem solutions) at the local and regional scale. This means less direct influence of national regulations on these solutions as long as some basic requirements like acceptable risks, groundwater protection, and a few others, are met. Robert Fowler (2007) noted common transitions in law and policy for contaminated sites in a number of industrialized countries. In Europe, for example, three generations in policy making can be observed. They will be described below. Some further remarks on the relation between the type of policy needed and the socio-economic context of contaminated site problems will be given in Section 23.8.

23.3.1 Generation 1: Command and Control Regulations by National Authorities

The early policies that arose around 1980 in a number of countries were a response to the discovery of a few severe contaminated site incidents that seem to require drastic approaches – at least that was the political and public perception in those days. This called for systematic inventories of contaminated sites, and sites were classified and prioritized using simple numerical criteria (concentrations) for contaminants in soil. Every site -small or large- was subject to the same assessment protocol, leading to relatively high costs for investigations of small sites, and incomplete information for large complex sites. Priority was based on the amount of contamination present at the site. The use of Risk Assessment was rather limited, because experience was lacking. Moreover the public did not trust Risk Assessment, which looked like a trick to allow contamination to remain in place. Solutions were either complete removal (dig and dump; pump and treat) of contamination or encapsulation. The regulations used to put these approaches into practice were simple top-down command and control regulations. These are very effective for those sites where human health or ecological risks are that serious that relatively fast action was required, irrespective of spatial planning or of other socio-economic considerations.

Due to the systematic inventories, many contaminated sites were identified that did not need immediate remediation, but that would require remediation if land use is changed into a more sensitive one. In those cases it is more appropriate to set priorities with regard to spatial planning rather then based on the degree of contamination. As the real “panic sites” were remediated or at least addressed, and many less urgent contaminated sites were identified, the public perception changed. Contaminated sites were no longer perceived as severe incidents requiring drastic governmental operations, but as a fact of life, especially in urban areas. The need for remediation and the priorities became more dependent on local special planning and a national policy giving priority to environmental impacts was generally felt as hampering spatial planning and socio-economic development. A policy change was needed.

23.3.2 Generation 2: Flexibility in National Regulations, Room for Local Site Specific Decisions

National policies were made more flexible as a response to the needs mentioned above. Spatial planning priorities were recognized. The room for site-specific risk-based fitness-for-use solutions increased. The closer link between contaminated site policies and spatial planning provided opportunities to use redevelopment as a driving force for improving soils and the built environment in general. Redevelopment of Brownfields (land previously used, for example for industrial purposes, which may be affected by contamination) and urban revitalisation initiatives led to the development of many new approaches. These approaches are more complex, but this can be an advantage, because more stakeholders are involved that can contribute to the solution and gain benefits from it, and different policy goals, like attractive urban landscapes and job creation, can be reached simultaneously. Many Brownfield redevelopment projects are now funded jointly by public and private parties. Public-private partnership financing became a preferred option and several variations on this theme were developed. However, remaining liabilities and possible future obligations for the land-owners after the remediation of their site are still major obstacles to Brownfield investment. Appropriate products of insurance companies can contribute to minimize risks and stimulate investments in Brownfields, but public authorities could also share or take over some of risks for future obligations and liabilities. This leads to a third generation of contaminated site management policies, which has just started in some countries, It should be noted that the third generation discussed below is not a replacement but an addition to the second generation in order to increase the success of this policy,

23.3.3 Generation 3: Regulations are Used to Create Opportunities and to Remove Barriers for Remediation by Private Parties

Central to the feasibility of remediation and redevelopment by private parties is their assessment of the social and economic viability of redevelopment of a site. This can already be quite complex in situations without soil contamination problems. The uncertainties concerning the amount of contaminated soil and groundwater present at a site, the risks (especially the perceived ones) for future use and functions, and uncertainties about the timeframe and the success of the remediation, may further increase the complexity and may form an additional barrier for redevelopment. There are three main categories of difficulties. One is related to costs and how these can be shared among public and private parties. This difficulty was already addressed in the generation 2 policy. The second difficulty is related to the risk of failure of the contaminated site remediation and how this risk is shared among the different parties. Potential failures translate into potential future costs and the sharing of these can be part of a public-private partnership as well. The third one is possible obligations and liabilities for contamination that can remain in place in view of the intended use of the site. Scientific insights about the risk of the contamination may change, and groundwater contamination may disperse more than expected, or future public authorities may want to change the use of the site. Landowners do not want to become problem-owners after redevelopment. Moreover most stakeholders and new landowners do not feel responsible for groundwater, which often extends beyond the boundaries of landownership. Regulations allowing public authorities to take over some of the risks for future liabilities and obligations for groundwater management and remediation are currently being explored. These regulations do not intend to weaken the polluter-pays-principle but to protect landowners from uncertain obligations and responsibilities and increase their willingness to cooperate as stakeholders in the redevelopment project. Due to this the voluntary remediation by private parties is likely to increase.

23.4 Contaminated Site Networks and Network Debates

The transitions between the three policy generations described above would have been much more difficult, if not impossible, without international exchange of information, experience and ideas. In Europe, for example, international networks and projects funded by EU research programs played an important role in shaping national contaminated site policies in the different EU Member States. Soon after the discovery of the first seriously contaminated sites – often called "hazardous waste sites" in those days – there was an urgent need to exchange experiences with experts from other countries on how to solve such new environmental problems. Initially, the improvement of remediation technologies was the main driver for international exchange. The earliest network for international cooperation, the NATO-CCMS (NATO Committee on the Challenges of Modern Society) group on contaminated sites, addressed in particular remediation technologies for contaminated soil and groundwater in a series of pilot studies (NATO-CCMS pilot studies 1985–2007). Contaminated site experts from science, policy and practice found their platform in the series of CONSOIL conferences (www.consoil.de). Policy makers started to network in the early 1990s when it became clear that contaminated sites are a structural problem in industrialised countries and not just a limited number of incidents. Networks like the worldwide operating "International ad hoc group on contaminated sites" (now called the "International Committee on Contaminated Land – ICCL) and the Common Forum (EU countries) arose and set the stage for improving contaminated site assessment procedures in the CARACAS concerted action (Ferguson et al. 1998) and risk-based site management procedures in the CLARINET concerted action (Vegter et al. 2002). Discussions with the industrial NICOLE network resulted in a common basis for Risk Assessment and contaminated site decision-making in EU countries, which integrates environmental and spatial planning related considerations. Some of the major themes in the network discussions are described below.

23.4.1 Environment Versus Spatial Planning as a Driver for Remediating Contaminated Sites

Traditionally, environmental policies often consider contaminated site problems from two main perspectives. The first is the perspective of protection – relating to the impact of contamination on human health and environmental quality. The other is the spatial planning perspective – managing the impact of contaminated sites on the way land is used, for example regenerating industrial areas, or increasing agricultural use, or for creating a nature reserve.

These different perspectives influence the different legal regimes used in different countries: some countries use environmental legislation as the primary means of preventing impacts from site contamination on land use and the environment, others use spatial planning legislation. In the extreme the environmentalist position often leads to the often costly obligation to restore the multifunctionality of the site, whereas the spatial planners position often leads to a fitness-for-current-use only approach, which may become a burden for future users of the site in case of a land use change.

The major conclusion concerning policy development identified in this discussion is to address environmental and spatial planning issues simultaneously. This became evident in the current holistic approach in urban development at contaminated sites. The holistic approach made it easier to link contaminated site remediation to economic issues, such as changes to land values and use of the market to drive environmental improvements.

23.4.2 Generic Soil "Numbers" Versus Site Specific Risk Assessment

The various approaches used to assess the significance of soil contamination and to set remediation goals in a number of industrialized countries has been heavily debated in the past. A review by Siegrist of this discussion was published as early as 1990. This report was the first comparison of different national policy approaches for contaminated sites and was followed up with reviews by Visser (1993), Ferguson (1999) and Carlon (2007). Siegrist noted that sometimes the actual remediation goal is not stated explicitly, but a decision on some acceptable course of action is reached on an ad hoc basis. If remediation goals are stated explicitly they are set by:

- Ad hoc site by site negotiation and decision-making.
- In reference to background levels.
- Application of predetermined Soil Quality Standards, guidelines and criteria.
- Site specific mathematical modelling, Risk Assessment and Risk Management decisions.
- Some combination of the options given above.

When only a few sites have to be dealt with, ad hoc approaches based on expert judgment and site-by-site negotiation and decision-making are appropriate. If the number of sites is larger, a programmatic approach was generally recommended. One should note, however, that within a programmatic approach, site-by-site negotiation and decision-making may still play an important role. The third generation policies as described earlier strengthen local decision making within flexible programmatic frameworks at the national level.

Visser (1993) gave the first detailed overview of policies in industrialized countries. This international comparison of approaches stimulated an intensive discussion of whether a programmatic approach should consist of predetermined Soil Quality Standards or more versatile site-specific modelling and Risk Assessment methods. Of course simple testing of measured concentrations in soil and groundwater against predetermined Soil Quality Standards is rather straightforward and by consequence less expensive than more elaborate Risk Assessment methods. However, remediation of contaminated sites is very expensive, especially for larger areas. Therefore, before deciding about the investment of large amounts of money in remediation, it might be expedient to invest in more advanced site-specific Risk Assessment. So a combined approach, using Soil Quality Standards to streamline the preliminary stages of an investigation and using site-specific exposure and Risk Assessment to achieve some fine-tuning in later stages, may be the most appropriate procedure. In fact the different approaches seem to converge to this combined approach as experience with remediation of contaminated sites increases, as was shown in the reviews by Ferguson (1999) and Carlon (2007). Current approaches in Europe and USA are generally framed as tiered assessment procedures. Generic Soil Quality Standards are used in the first tier as a "quick and dirty" preliminary assessment tool. In second and higher tiers site-specific Risk Assessment is used which involves modelling and measuring the actual exposure of targets of concern at the site. Because generic Soil Quality Standards are just a tool in complex assessment procedures and can be formulated in different ways in different countries, a direct comparison of the numerical values of Soil Quality Standards without considering the complete assessment procedure and the underlying assumptions would be misleading. Moreover the assessment procedure and the derivation of Soil Quality Standards may also depend on country-specific political choices and geographical elements. The study of Carlon (2007) took geographical and political elements into account in order to make a comparison as realistic as possible.

The use of various Risk Assessment approaches and their scientific background have been discussed in the CARACAS network. CARACAS observed that Risk Assessment is not a fully-fledged scientific discipline on its own. It is a rather loose assemblage of elements borrowed from various scientific disciplines. There is a large need for further improvement and integration. Constructive discussions with the NICOLE network (Network for Industrially Contaminated Land in Europe) about the uncertainties in the assessment and the level of precaution to be applied in view of the current status of assessment procedures led to the common position that Risk Assessment is already a very useful tool, if one is aware of its limitations and uncertainties.

23.4.3 Risk Management

Until quite recent times, the remediation of contaminated sites was based on civil engineering approaches aiming at maximum risk control (excavation or containment). These approaches are still the fastest way to solve the problem. In a densely populated country like the Netherlands there is heavy pressure on re-utilization of contaminated sites, so fast remediation is an advantage. However, the high costs associated with these methods are also a big disadvantage and were prohibitive in many redevelopment projects. This has led to the exploration of risk-based solutions which make use of natural capacities of the soil environment like immobilisation, isolation and natural attenuation of contaminants. Remediation methods that make use of natural capacities of the soil environment –or expressed in a more modern way- ecological services of soil, natural attenuation, more specifically Monitored

Natural Attenuation (MNA), is nowadays widely advocated as a low cost solution for soil remediation (see Chapter 22 by Peter et al., this book). It involves degradation of contamination by natural processes. These methods are less costly. But they require more time, have less predictable results and may require long time monitoring and aftercare. This is a drawback if land can only be reused after a lengthy remediation.

A possibility to combine fast-acting measures with a longer term extensive treatments becomes interesting in these situations, since it may provide an optimal balance of Risk Management, maximising wider environmental merit and limiting costs. Moreover, the soil itself has some interesting characteristics, which may help in reducing the risk. Soil has a natural capacity to act as a barrier, which can be used in containment approaches and it has a natural capacity to biodegrade contaminants. If these natural capacities can be used, the costs of solutions as well as unwanted side effects like CO_2-emissions through energy consumption will decrease. The use of the natural capacities of soils in contaminated site management solutions need to be further explored both from a scientific and a regulatory point of view, in order to meet the general sustainability requirements of soil protection.

Whether natural attenuation is actually successful will depend on site-specific conditions. These are never completely known, and processes like biodegradation and chemical and physical behaviour of contaminants may be difficult to predict. Natural attenuation may better be considered as an experiment under controlled conditions. The authorities will have to provide some room for these experiments in terms of time and space. The control of these experiments will require monitoring of the process. At certain points in time one must be able to decide whether natural attenuation is performing as expected or whether one has to resort to more intensive Risk Management options.

Natural attenuation may also be enhanced by addition of contaminants to soil and groundwater in order to stimulate microbial activity. If the presence of these contaminants and residual contamination is temporal or has negligible effects on the quality of the groundwater, this may be considered acceptable from a soil and groundwater protection point of view. It is up to the policymakers to find the right balance between soil remediation and soil protection in these cases, but it is clear that MNA solutions require longer-term management by the competent authorities. The network discussions provided valuable ideas how this management approach should be developed.

23.5 A Policy Maker's View on Risk Assessment for Contaminated Sites

Although “Risk” is becoming a central concept in environmental policies, this does not mean that risk is an easy and well-defined concept. Definitions do abound, at least 20 definitions ranging from informal to very formal (mathematical) were mentioned by Vlek (1990). Risk Assessment, risk analysis, risk perception, policy-making and decision-making are being quite extensively studied in the social sciences and psychology. Although relevant for the development of Risk Assessment procedures and Risk Management decisions for contaminated sites, most studies about “the nature of risk” do not address contaminated sites specifically. From a natural scientific point of view, the current approaches in Risk Assessment for contaminated sites are still rather loose assemblages of concepts and methods borrowed from various disciplines (Ferguson et al. 1998). Nevertheless, Risk Assessment is considered a very useful tool in environmental policy since it provides a rational and objective basis for priority setting and decision-making. The use of Risk Assessment in contaminated site problems is nowadays advocated by both regulators and industries.

23.5.1 A General Framework for Risk Assessment for Contaminated Sites

At a general level, most EU countries have a common framework for Risk Assessment procedures (Ferguson et al. 1998, Ferguson and Kasamas 1999). Generally, the following protection targets are considered:

- *Human health*.
- *Ecological functions*.
- *Groundwater:* This risk is related to the dispersion of contamination in groundwater. Criteria do vary among countries and so do protection levels. In some countries, groundwater is protected as a resource that has to remain pure. Other countries will use risk-based protection levels.
- *Construction:* Effects of soil contamination on structures and construction works is explicitly mentioned in some countries. In other countries, this endpoint may be more implicit.

Risk Assessment procedures focused on the two major protection targets (human health and the soil ecosystem) have been described in detail in this book, in “Humans Health Aspects (Part III)” (Chapters 5, 6, 7, 8, 9, 10, 11, and 12) and “Ecological Aspects (Part IV)” (Chapters 13, 14, 15, and 16). Risks related to groundwater migration have been described in “Groundwater-Related Aspects (Part V)” (Chapters 17, 18, and 19).

A general view of the different steps to be taken in Risk Assessment, from suspicion to preliminary investigations and to conclusions about risks and communication is illustrated in Fig. 23.3. Risk Assessment of contaminated sites usually starts with some suspicion about facts that may have led to the presence of soil contamination: handling or storage of relevant amounts of hazardous substances on the site, accidents, spills etc. The next step often is a historical “recherche”; information is gathered, which is available in archives and other accessible sources. This is usually accompanied by a visit of the site. If these steps strengthen the suspicion, a further investigation may be carried out and results may be compared with generic Soil and Groundwater Quality Standards or can be evaluated using models to predict adverse

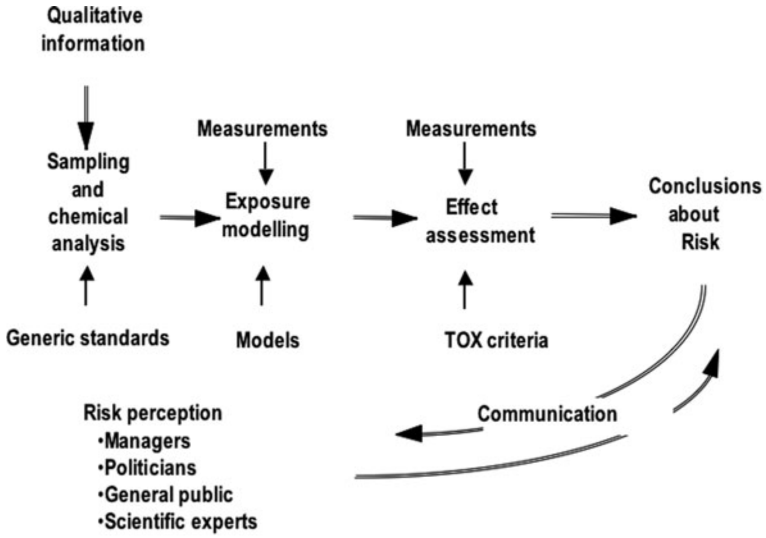


Fig. 23.3 A general view of the different steps to be taken in Risk Assessment

effects of contaminants and the need for remediation. These model conclusions can be substantiated by actual measurements of exposure or even epidemiological data. In theory the following assessments are possible:

- Qualitative information about the history of a site which is assessed by experts.
- Results from chemical analyses of soil and groundwater samples are compared with soil and Groundwater Quality Standards.
- Results from chemical analyses of soil and groundwater samples are used to compute exposure with exposure models. Exposure is than compared with toxicological references to allow conclusion about risk for human health or ecological functions.
- Exposure can be measured instead of estimated with models. Results from these measurements can be compared with toxicological references to allow conclusion about risk for human health or ecological functions.
- Measurement of exposure can also be used as input to a toxicological dose response model.
- Adverse effects can be assessed by epidemiological studies at the site.

In general, the less prediction the more reliable is the assessment, but the more difficult is the investigation. This trade-off could lead to a system of assessment methods with various levels of sophistication. However one should be aware that the traditional investigation protocols that have been adopted in various national contaminated site policies do not sufficiently recognise the trade-off between cost of investigation and the level of detailed information needed to make a decision. Usually a fixed sequence of investigations seems to be mandatory. Historical information is used to classify a site as "suspect", and a first investigation measuring concentration of contaminants is to follow soon. If contaminants are above a trigger level, site-specific Risk Assessment is performed, usually involving models. If the site is a risk for human health or the environment in view of its current or planned new use, an investigation will follow to assess the remediation options.

In some cases, however, it is already known on the basis of historical information and experience that a site will need remediation. If there is enough certainty to construct a conceptual model of the site, that adequately describes the intensity of the sources and pathways that lead to exposure of the receptors of concern, one may skip the time-consuming and costly site specific Risk Assessments and proceed to the remediation phase. It should be noted that juridical requirements concerning the validity of remediation decisions may lead to the use of site specific Risk Assessments in all cases.

At a general level, the framework for Ecological Risk Assessment may be very similar to the framework for the assessment of human health risk. The target of the assessment, which may be defined as "Ecosystem Health", is however more complex than human health, since it involves a large number of phenomena operating at various spatial and temporal scales. As far as the ecological risks of contaminated sites are concerned, it may be important to make a distinction between two types of ecological risks: those occurring at the contaminated site itself, and the impact of the site on the surroundings, either by transport of contamination or, for instance, by the loss of an important habitat within a limited range of habitats of plant or animal species. The first category is the effect of the contamination on the capacity of the soil to support plant and animal life at the site. The impact of the site on the surroundings may be addressed by procedures quite similar to an environmental impact assessment. Instead of assessing the environmental impact of potential contamination from, for instance, a new industrial site, the impact of the actual contamination is assessed for a site in its current state. There is, however, also a large difference between environmental impact assessment and contaminated site Risk Assessment. Environmental impact assessment is used to choose between locations and preventive approaches, whereas contaminated site Risk Assessment addresses the actual contamination resulting from "choices" of the past.

23.5.2 Risk Assessment and Risk Management

The use of Risk Assessment in environmental policy probably originated in the USA. The aim of the methodology was setting priorities for environmental protection in an objective and scientific way, avoiding political "perceptions" of the administration (National Academy of Sciences; (NAS 1983)). The distinction between Risk Assessment (the objective scientific part) and Risk Management (the policy driven decisions about risks) was of utmost importance in these early ideas. Recent discussions, however, seem to question the strict separation of assessment from management for a number of reasons:

(1) Politicians have to consider the perception of non-experts. The perception of the public might frame the questions to be asked during an assessment. A formal scientific assessment may not fully answer the political questions.

(2) If scientific assessments are made very detailed and very specific, they almost dictate the decision to the risk manager. On the one hand this simplifies and accelerates the decision-making, while on the other hand the autonomous role of the decision-maker is reduced. This trade-off calls for a dynamic interplay between assessment and management of risk.

(3) Whether a certain risk is acceptable is not a scientific issue but an individual choice, if it concerns a single person, or a political setting if a large group of people is concerned. Scientists can assess the likelihood of occurrence of an adverse effect. However, if they act according to the scientific tradition that there is no effect unless it is proven beyond all statistical doubt, Risk Management will not be very protective. Risk Assessment will, on the other hand, fail to come to a decision, if every kind of effect shall be excluded. According to Shrader Frechette (1996), scientists involved in Risk Assessment must avoid false positive conclusions about risk as well as false negative ones.

    Risk Assessment usually should result in a decision: "risk" or "no risk". In many cases however, the decision between these two opposite endpoints is difficult and costly. The assessor experiences, often for a long time, a situation where a risk cannot be proved to be present or to be absent with a high degree of probability. The less doubt is desired about the decision, the more effort is necessary. This leads to another question in Risk Assessment: How much doubt is tolerable?

(4) Risk Assessments can involve statements about complicated and poorly understood phenomena. Even for situations with adequate dose-effect relationships, a no-effect level (or negligible risk level for carcinogens) may be estimated, but cannot be verified in practice. To detect low levels of risk, very large sample sizes are needed. Scientific proof is possible in theory, but not in practice. This problem has been labelled as "trans-scientific" by Weinberg (1972).

If a Risk Assessment is hampered by lack of information, large uncertainties and numerous trans-scientific questions, it may not be very powerful in helping to make decisions. Some people might claim: it is the best we have, because it is the only way to make an objective decision. Others are more sceptical: wrong decisions are wrong even if objective! It is clear that Risk Assessment cannot solve all problems in decision-making in uncertain situations, which may lead to provocative statements like O'Brien (1994): "Scientists should move away from the hubris of assimilative capacity estimation and Risk Assessment to the wisdom of a precautionary orientation". One might also say that if scientific Risk Assessment is hardly possible, a more intuitive political perception of risk (like the precautionary principle) might provide better guidance in decision-making. Therefore, a modern view on Risk Assessment does not separate Risk Assessment from Risk Management. Risk

Assessment is nowadays seen as an important management tool to make decisions about environmental problems with many uncertainties.

Part VI in this book includes information on the theory and application of Risk Management procedures.

23.5.3 The Role of a Scientist in Risk Assessments

The scientific definition of risk as "a combination of the consequence of a negative effect and the probability of that effect to occur", and the large uncertainties in risk estimation, both encourage the use of statistical approaches and the application of decision rules based on the amount of uncertainty. There are several types of uncertainty to be dealt with in a Risk Assessment. A number of authors have tried to classify them. Wynne (1972), see also Doves and Handmer (1995) for instance presents the following taxonomy of uncertainty:

- *Risk*: system behaviour is known, and outcomes can be assigned probabilistic values.
- *Uncertainty*: important system parameters are known, but not the probability distributions.
- *Ignorance*: What is not known is not known; and the degree increases when the level of action or commitment-based on what we think we know increases.
- *Indeterminacy:* causal chains, networks or processes are open, and thus defy prediction.

Shrader Frechette (1996) suggests a number of rules for scientists involved in Risk Assessments. In Risk Assessment and environmental decisions four classes of uncertainty are considered most relevant. The four classes are:

23.5.3.1 Framing Uncertainty

This type of uncertainty is related to the translation of the policy question in a scientific question. Do we have to prove beyond reasonable doubt that there is a risk, or that there is no risk? In the first case lack of information seems to promote safety, in the latter case it increases the risk. It is necessary to use a so-called three valued frame: "there is no risk", or "the decision is not possible due to lack of information", or "there is a risk".

23.5.3.2 Modelling Uncertainty

This type of uncertainty pertains to the realism of models, and to the question of the reliability of model predictions. Very often models are considered to be validated or verified if the output of the model is consistent with some other model. The only valid test is the comparison of model predictions with real world data. If real world phenomena are successfully predicted one might gain confidence in the model, but asserting that the model is “true” on the basis of successful predictions is scientifically speaking not possible. In science, empirical data can only lead to falsification of hypotheses and not to their confirmation.

23.5.3.3 Statistical Uncertainty

This type refers to the so-called type I and type II errors in statistical approaches. Scientists tend by tradition to minimize type I errors, the risk of rejecting a true hypothesis. Type II, the risk of accepting a false one, may be more serious in environmental problems, especially when the hypothesis was that there are no human health risks or ecological risks.

23.5.3.4 Decision Theoretic Uncertainty

This type of uncertainty shows up in Risk Management decisions. It relates to the following questions. Should the worst-case scenario govern the decision even if it has a very low probability of occurrence? Or must the decision be based on the more likely scenarios? Or should decisions be based on utilitarian principles and cost benefit analyses?

23.5.4 Risk Perception and Communication

It is well known that the results of the formal natural scientific approach in Risk Assessment can be very different from the risks as perceived by the public. Risk perception may be governed by a number of factors (Vlek 1995), but the main difference is that risk is perceived more intuitively.

The factors influencing risk perception and acceptance are important for communication of results of formal Risk Assessment to the public. However effective communication does not exclusively depend on the perception of the public but also on the characteristics of the communicator, the message and the medium that is used. It is beyond the scope of this paper to give a full treatment of communication strategies for management of contaminated sites. Much can be learned from experience from Brownfield redevelopment (see Chapter 25 by Nathanail, this book), where risk communication becomes one of the many aspects of an overarching public consultation and stakeholder engagement strategy (EUBRA 2007). One final suggestion related to Risk Assessment is that it may be easier to communicate about soil quality (this piece of land is still fit for . . . and will be fit for . . . after remediation) than to communicate about risk (there is still contamination left but the authorities state that your risk is acceptable).

23.6 Risk-Based Land Management – The Concept

Policy makers and regulators, as well as other stakeholders, need to make balanced and informed decisions about contaminated sites. To stimulate this cooperation, the CLARINET concerted action developed the Risk-Based Land Management

(RBLM) concept. It is a framework for development of policy, research and practice in sustainable management of contaminated sites. It focuses on historical contamination (contamination resulting from past practices) and allows for regional and site-specific solutions.

The RBLM framework structures the decision-making process to achieve sustainable solutions. It integrates two key decisions for the remediation of contaminated sites:

1. The time frame: this requires an assessment of risks and priorities, but also consideration of the longer-term effects of particular choices.
2. The choice of solution: this requires an assessment of overall benefits, costs and environmental side effects, value and circumstances of the site, community views and other issues.

These two decisions have to take place at both an individual site level and at a strategic level, especially as the impact of contaminated sites on the environment cannot have a large scale regional dimension only, but also potentially wide-ranging long-term impacts. The decision-making process needs to consider three main components, which form the core of the RBLM concept:

1. Fitness for use.
2. Protection of the environment.
3. Long-term care.

The aim of the RBLM concept is to achieve the integration of approaches originating from different perspectives (for example spatial planning, environmental protection, engineering), based on the identification of common goals:

- optimised use and development of technical and administrative solutions;
- sustainability – evaluating and optimising environmental, economic and social factors.

23.6.1 The Term "Risk-Based Land Management"

The term Risk-Based Land Management appears similar to other expressions used in the context of soil contamination, for example Risk-Based Site Management. However, RBLM considers the issues from a larger scale perspective, and covers the full range of contaminated site problems for which regulators and decision makers are responsible. The constituent terms of the concept are carefully chosen and are used as follows:

23.6.1.1 Risk

Risk describes the combination of the probability and the effects of contamination, for example adverse effects on human health, on ecosystems or on water resources. Adverse environmental effects are generally considered as the result of a process where some potential hazard (a toxic chemical or other agent) affects a target to be protected (people, animals, plants, ecosystem processes, water resources or buildings). For this process to operate, there must be a connection (a pathway) between the potential hazard (the source) and target for protection (the receptor). If an adverse effect has occurred, the consequences are often described as damage. Poor soil and water quality may in turn lead to adverse social and economic effects.

23.6.1.2 Land

Land represents a geographical area. For example, it could be a single site, a number of sites in the same locality, or it could be a region such as municipality or larger area. It also includes the soil, surface water and groundwater beneath the surface of the sites, adding a third dimension to the traditional spatial planning interpretation of land. The reason why the word land is used here instead of sites is that RBLM wants to put contaminated sites in a spatial planning context and to move away from the classical management of incidents approach for individual contaminated sites.

23.6.1.3 Management

Management is a set of activities involving decisions about assessment, remediation, land use restrictions, monitoring, spatial planning, aftercare and other issues. In the context of Risk Management, it is a much broader activity than the selection of a remediation technique. It includes all aspects of developing and implementing a sustainable approach. The scope of this management may also be wider than the contaminated sites issue. Other environmental impacts and stressors may need to be dealt with at the same time. There are also different types of "managers" of land – they may be, for example, the owner or user of an industrial site or a municipal authority.

23.6.2 The Components of Risk-Based Land Management

23.6.2.1 Fitness for Use

Fitness for use relates to the reduction of risks to human health and the environment as much as necessary to ensure the safe use or reuse of the site. It focuses on quality requirements of the site for the land use and soil functions, and takes into account the timeframe of the particular use of the site. For example, the assessment considers how long a receptor might be exposed to contamination.

Risks related to the use of the site should be "acceptable" for the people concerned. This acceptance might be obtained if the quality of the site meets certain minimum quality requirements. In some cases, obtaining acceptance might require additional quality requirements to create confidence and security. It is essential in determining the "total quality requirements" to know all the aspects of the site use.

This will ensure that the requirements are appropriate. It is also necessary to consider the future activities and controls on the site with the purpose of ensuring that long-term risks are also managed, and that the site will continue to be "fit for use" in the future.

Making certain choices about the management of the site can not only achieve the necessary quality requirements in relation to immediate fitness for use but also improve the quality of the site over time. For example, introducing additional gradual treatment would open up opportunities for land use changes, an increased Biodiversity and less long-term care.

23.6.2.2 Protection of the Environment

Protection of the environment is related to the wider effects, in contrast to those only related to the use of the site. In the UK the term “suitable for use” combines the two concepts of fitness for use and protection of the environment (UK/DETR Circular 2/2000). Protection of the environment has two objectives:

1. To prevent or reduce negative impact on the (natural) surroundings, including Ecosystem Health and Biodiversity, groundwater bodies, surface waters.
2. To conserve and, if possible, enhance the quality and quantity of resources (for example land, soil, water or cultural heritage).

The requirement to achieve both fitness for use and protection of the environment means that solutions have to be chosen carefully. A solution that meets only the fitness for use requirements is probably not the best solution if it creates potential problems in surrounding areas. A solution that manages the risk of contaminant dispersion may be different from the solution that manages risks to achieve “fitness for use”.

Solutions may in turn lead to the exploitation of other resources, such as energy reserves, or land capacity for disposal. Other environmental and spatial planning policies will aim to protect these resources and a balanced decision – or new solutions – will be needed where there is conflict between the objectives of risk reduction and conservation of resources.

The decision to conserve land or soil as a resource may lead to policies favouring redevelopment of Brownfields over Greenfields (non urban land). This in turn may lead to increased pressure to develop new solutions to deal with the risks to human health and the environment. It also shows the need for strategies to prevent sites from becoming Brownfields.

23.6.2.3 Long-Term Care

If a solution leaves contamination in the soil, there is a need for long-term care. Monitoring and control may be necessary to ensure that the solution remains appropriate, that it continues to work and that any restrictions on future choices regarding the land use are enforced.

Solutions that are based on the current land use only, or rely on specific restrictions on land use, need additional documentary records. Taking into account the social and economic burden of long-term care and the risk of failure is essential in identifying sustainable solutions.

23.7 Application of RBLM in Practice

The way in which the balance between the three components of RBLM is achieved will be different for different remediation and Risk Management approaches. Over the past fifteen years, developments in contaminated site policies and the emergence of a wide range of treatment approaches have broadened the repertoire of potential solutions for contaminated site problems. There can be other options rather than only "dig and dump" or containment. However, it is clear that there is no universally most practical solution. Each solution has its advantages and disadvantages, which depend on a wide range of factors and requirements, such as:

- nature of the contamination;
- physical characteristics of the site;
- use of the site, either current or planned;
- environmental setting, in particular ecosystems and buildings;
- hydrogeological characteristics and impact on water resources;
- nature of the impact on the community;
- local and regional practical feasibilities.

The choice of any specific practical solution, either at a strategic policy level or for a particular site, needs to take into account the extent to which the site meets any fitness for use criteria, achieves adequate protection of the environment or needs longer-term care. This assessment is complex, and has generated a demand for decision support tools, which may vary from straightforward information about the broad advantages and disadvantages of various options to formalised complex weighting systems. The issues that the risk-based site manager has to address in order to ensure a sustainable solution include:

1. Risk reduction.
2. Land use related requirements.
3. Using natural capacities in the soil and water environment (see Section 23.4).
4. Costs.
5. Involving stakeholders.
6. Managing uncertainties.
7. Other management constraints and influences.

To assist decision-making these issues are discussed below based on the underlying three components of Risk-Based Land Management: fitness for use, protection of the environment and long-term care.

23.7.1 Risk Reduction

The initiation of decisions relating to risk reduction comes from the results of Risk Assessment, which in turn can inform the choice and objectives for solutions. It is important to note that the importance of human health risks, ecological risks or more general environmental risk like the risk of spreading of contamination in groundwater, surface water or air may receive equal weight in the Risk Assessment of the site but in the design of the solution things may look different, dependent on the requirements for future use. Human health and ecological requirements for parking lots are of course much less stringent than for houses with gardens where human exposure is more likely and ecological functions are needed for the gardens. The risk of spreading of contamination may dominate the discussion about risk reduction in the parking lot situation. The final choice about risk reduction needs to consider the key decisions within RBLM referring to the time frame and the choice of the solution.

23.7.1.1 The Time Frame

One key issue in relation to the timeframe is determining when to intervene, control, or reduce risk. This requires prioritisation and Risk Assessment procedures based on an assessment of the fitness for use of the site and of the need for protection of the environment. Many Member States of the EU and industries have already developed, or are in the process of developing, this type of approach in their decision-making frameworks for contaminated sites.

The CARACAS network considered a number of aspects of Risk Assessment relevant to determining when action is necessary and to priority setting in relation to human health and environmental risk like risk for ecosystems, spreading of contaminants in groundwater or risks for specific ecological functions. Clearly, in terms of total resource use, it is equally important to decide when it is safe to postpone action on sites to avoid unnecessary expenditure.

However, the need to conserve or enhance a resource – for example soil or water – may also set the timeframe for action. Action may be justified rather sooner than later, because the long-term care requirements are too great a burden. There are other drivers for action, for example the need to reuse the site or to increase the value of the land for the owner or developer. The urgency of a site remediation is therefore a function of one or more of these environmental, economic and social factors, and for most large projects a combination of all of them.

23.7.1.2 Choice of Solution

The factors that affect the urgency or priorities for action also determine a key interface between timeframe and choice of solution, which is when the solution needs to become effective and for how long. Another factor in the choice of solution is the way in which the solution relates to the underlying risk. The process underlying the

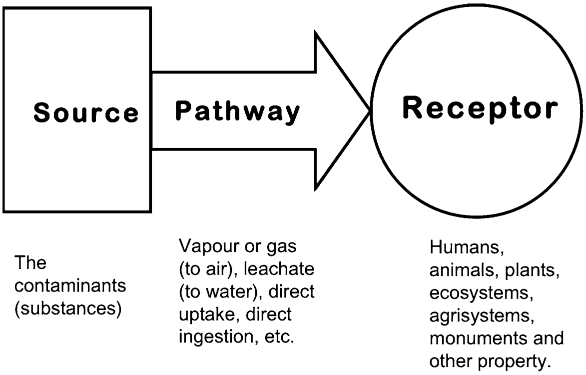


Fig. 23.4 The source-pathway-receptor causal chain

mechanism of environmental risk is often described in terms of a source-pathway-receptor causal chain (Fig. 23.4). This is one of the fundamental approaches to the assessment of contaminated sites and development of solutions. In this context, the source is the contamination in the soil or water and the pathway is the link of the contamination and the effects at the receptors of concern.

In principle, risk reduction may be achieved by removing the source, by controlling or eliminating the pathway, or by removing the receptors. Each of these different approaches has a different interface with the three components of RBLM. For example, at the simplest level of analysis, removing the source may use other resources and have an overall negative impact on the environment. Control of pathways may have a lower initial resource use, but has long-term care implications. Removing receptors may be costly or socially problematic.

23.7.2 Land Use Related Requirements

23.7.2.1 Practical Needs

Different land uses have different needs in terms of the condition of the site. For example, some land uses require direct access to the soil, preventing the use of surface containment measures like capping with concrete or asphalt. Others may require the preparation of the site for geotechnical purposes, for example to support foundations. This type of considerations can be combined with the “fitness for use” principle, resulting in an assessment of “fitness for purpose”.

In some cases, the contaminated layers in a site may contain rubble, rubbish and coarse waste materials requiring excavation for other reasons, such as construction work. This could make excavation and removal an efficient solution to risk reduction. Often on-site recycling and re-use of debris is possible and can reduce demand on primary aggregate resources in these cases.

23.7.2.2 Spatial Planning Requirements

Spatial planning decisions involving contaminated sites cannot be made effectively without considering the quality of the soil and the nature of any site management solutions used to deal with contaminated sites. At a site-specific level, the most important decision is usually the quality of land for a particular use, whether this is for the current use, or in relation to potential future uses. Even where there is a specific use, some of the particular details of use may not be realised at the time of treatment of the site. They may include unmonitored changes, where, for example, an owner carries out activities allowed as part of his normal use of the site. A typical example is where the permission to build houses also allows later additional work by the house occupier, for example to build an enlargement of the building which may involve soil excavation. Where there is a known use, the assessment of fitness for use should consider all possible additional activities on the site allowed within the use classification.

It is also essential to consider the effect of the use of the site, or a change of use of the site, on the characteristics of the site – including the hydrogeological characteristics – and the behaviour of contaminants. This may not only affect the fitness for use of the site, but also may cause the contamination to have an unacceptable impact on the environment. This may make the particular use of the site an impractical proposition.

There may be a very different land use in the future. For example in Brownfield redevelopment where commercial or even industrial areas are being converted to more sensitive land uses such as housing or leisure activities. Long-term management of information is essential to retain knowledge about what has been done and what future implications there might be for the land use. This may include record keeping such as land “log-books” or registers, including information on site condition.

A land use change in the future requires a detailed consideration in relation to the components of RBLM. The standard of remediation that leaves the site fit for a specific use would need reassessment for a new land use. If treatment of the contamination has been based on controls of exposure pathways (or on other ways in which the receptor has access to the contamination), or if the source reduction has led to a limited overall standard, then the land resource may not be as flexible. For the future, at a strategic level, this may not be adequate in conserving land as a resource.

Incorporating RBLM into spatial planning policies and systems is one way of managing contaminated sites effectively. Strategic policies can consider the overall interaction of development with land quality to achieve optimum use of land. This includes considering the wider issues of environmental protection and resource conservation. It may not be appropriate to allow certain types of development, for example, if they will cause an unacceptable environmental impact, e.g., by mobilising contaminants, or restricting or interfering with other environmental improvements like groundwater remediation. On a site-specific basis, spatial planning controls can manage the continuing implications of long-term care requirements linked to particular solutions.

23.7.3 Using Natural Capacities in the Soil and Water Environment

Within the RBLM framework it becomes interesting to explore cost-effective remediation options, using a conceptual model as illustrated in Fig. 23.5.

In particular the question of how remediation approaches that use natural capacities of soils or groundwater, like natural attenuation (see Chapter 22 by Peter et al., this book), can meet the general requirements for sustainability. For example, the use of the soil for containment will have to meet the same requirements as other containment techniques, such as waste disposal sites. These requirements will relate to the need for the site to be fit for use, for sufficient protection of the environment (including resource use) and for long-term care. One should also be aware that sites may become commercially less attractive if long-term natural attenuation is used as remediation, or immobile contamination remains in the soil. Many developers seem to prefer land without use restrictions or obligations of long term care.

Another relevant issue is the extent to which the resource "soil" or "groundwater" is considered as a resource to be protected as a receptor (irrespective of its use) and not just as a pathway to other receptors, like sources for drinking water. In the EU for example, the Groundwater Directive (2006/118/EC) prohibits any discharges of hazardous contaminants into the groundwater. This may restrict the use of contaminants that enhance biodegradation. However, if the application of certain substances in groundwater to stimulate natural attenuation is temporary, of short persistence and has negligible long-term effects on the quality of the groundwater, it could be

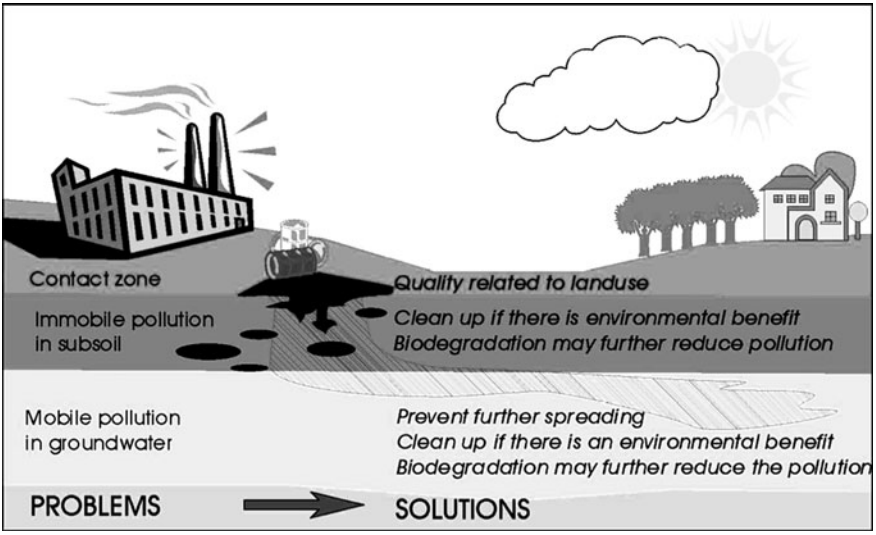


Fig. 23.5 Exploiting the natural capacities of the soil and groundwater system in risk-based site management

classified as acceptable scientifically – even from a long-term point of view – for soil and groundwater protection. The same may hold for small-scale dilution and dispersion of the polluting substance when it degrades in the soil matrix.

According to the European Union's viewpoint it is vital that soil and water resources are valued and protected properly. There are two equally important goals for integration of resource protection with contaminated site management. One is to prepare an optimal plan combining use of the resource with its protection and restoration. The other is to consider all the components of RBLM for each site to ensure the full assessment of implications for natural resource protection, conservation and enhancement. Considering natural processes and capacities in the soil and water environment allows for a wider range of potentially sustainable solutions.

23.7.4 Costs

23.7.4.1 Types of Cost

The cost of remediation work is often an overriding factor in decisions. It is all too easy to take a narrow view of costs based only on obvious direct costs of remediation and to ignore wider types of costs, some of which are related to the components of the RBLM analysis, such as:

1. Indirect or "opportunity costs", for example where the site is not fit for use because remediation has not taken place, or where there is a loss of income caused by delay in being able to use the site because the remediation process is slow, or where a particular land use is not possible because of a restricted remediation.
2. Components related to protection of the environment, for example to prevent environmental damage or burdens falling on different sectors or "third parties", or when a remediation process which uses a large quantity of off-site material may in effect place the burden of renewing that resource onto someone else, or a negative impact on an ecosystem may not be noticed for many years.
3. Long-term components, particularly where there is a requirement for long-term care, or where a resource may need to be restored in the longer-term.
4. Costs relating to different financial mechanisms for providing the funding, such as those which use income and those which require transfer of capital or exchange with other assets.

The comparison of costs can be made more complicated by financial accounting conventions and by the tendency to avoid financial costs today, when they can be postponed to the future.

23.7.4.2 Balancing Costs and Benefits

Costs and benefits can be considered within RBLM as part of integrated decisions considering fitness for use, protection of the environment and long-term care. In particular, RBLM offers a structured way of identifying the benefits and drawbacks of different options. These can then be balanced against the costs to produce a sustainable solution. RBLM will also make cost comparisons more transparent.

23.7.5 Involving Stakeholders

Contaminated sites are the same as other environmental issues in terms of the range of "stakeholders" – those who have a direct or indirect interest in the outcome of decisions. For some aspects it is more complicated, since it often touches at the heart of a society or individuals in that it affects not only their own immediate environment, but also the value of something precious to them: their land.

Dialogue with stakeholders may affect the choice of certain solutions over others. It will have to deal with "values" which are difficult to express in terms of risk or utilitarian concepts like land use or soil function. The conservation of a pristine underground environment and the conservation of geologically or archaeologically important sites are examples of this.

23.7.6 Managing Uncertainties

There are many scientific and technical uncertainties in contaminated site decision-making (Ferguson et al. 1998). Uncertainties will always be there. RBLM provides a way to deal with them in a systematic and explicit way. This is important, because in Risk Management, some approaches lead to more certainty than others.

23.7.6.1 Technical and Scientific Uncertainties

For instance, treatment to reduce contaminant availability is more likely to lead to management uncertainties than excavation of contaminated soil. If the remediation aims to reduce a particular degree of risk, there will be uncertainties in the scientific calculations of the remediation goal. Knowledge about the toxicity of contaminants may change in the future, leading to either more strict or less strict remediation targets. In addition, there may be contaminants present in the environment, which have not yet been identified as potentially toxic.

Scientists may discover much more in the future about the way in which natural resources respond to pressures such as contamination, which could lead to a greater or lesser need to protect them, or a need to manage the remediation of contamination in a different way to work with these responses. Inevitably, there will be aspects not yet realised that have implications for long-term care, for example the effect of climate change on other soil and groundwater behaviour that in turn affects contamination.

Apart from the scientific uncertainties, there may be uncertainties in the needs of society, particularly in the future. This is not just a matter of knowing whether the land use may change or not. The way sites are used in the future may be very different from the types of land uses known today. For instance, it may be necessary to consider whether land use related remediation, particularly when based on control of the pathway or receptor, may prevent future use of sites for underground building and infrastructure.

23.7.6.2 Decision-Making

These uncertainties translate into site management problems, both with regard to environment and to spatial planning. However, managing scientific uncertainty is not a new concept. One of the most important political choices concerning decision making in uncertain situations is the degree of conservatism of precaution. There are two extreme positions:

(1) to set very stringent and highly precautionary requirements (building in very conservative safety factors to allow for unknown scientific developments) or
(2) to take a *laissez faire* attitude to assessing the risk (assuming that problems caused by lack of action today will be dealt with by the society of tomorrow, as today's society has to deal with yesterday's legacy)

Most contaminated sites management approaches fall between both extremes described above.

To put policy into practice, one must also make choices between simple decision support tools or more sophisticated ones. Decision support tools like generic remediation objectives (standard values for concentrations of contaminants) for instance are easy to apply, but many stakeholders feel that the use of generic soil quality standards leads to arbitrary decision-making. Generic soil quality standards do not reflect the site-specific risk of contamination in most cases. On the other hand, introducing time-consuming, costly and complex site-specific procedures that do not easily allow decisions when factors are unknown for even the most simple contaminated site problem, is not very expedient. In general some intermediate tool between the two extremes described above would be adequate.

23.7.7 Other Management Constraints and Influences

Whatever the outcome of the theoretical analysis of the components of RBLM is, there are other external factors influencing its application in practice. One set of factors relates to the decision-making process. Who decides and how do they decide? Will it be a dynamic and open decision-making process, involving all interest groups, or can a single decision-maker apply a protocol or a mandatory decision support system?

The conceptual idea of a "manager" in the RBLM approach does not automatically imply that there is a single decision-maker. The "manager" may be the competent national, regional or municipal authority. These authorities will have to act within their mandate to represent public interest. For industrial sites, the manager may be the owner, who can make decisions within certain limits, imposed by the authorities. Individual site managers may have other boundaries set by corporate policies.

Clearly, whoever the decision-maker is, good decisions are based on good information, on understanding of both the short-term and long-term consequences and on a systematic approach. Transparency as well as recording of assumptions and uncertainties is also essential to promote confidence and manage the site in the long-term if the solutions for the site require long-term care.

Another set of factors influencing the application in practice relate to external constraints of the process. There may be requirements set by those providing the funding, whether this is direct (for example from public authorities or from investment by others) or whether it is indirect (for example from the citizen in the form of taxation). Legal constraints may prohibit specific treatment and Risk Management solutions. Some of these constraints may be essential interfaces with other aspects of protection of society and the environment, others may require careful application and review if they endanger a fully sustainable solution for contaminated site problems.

23.8 Concluding Remarks

The development of contaminated site policy frameworks as described in this chapter can be seen as a sequence of three generations. Although the three generations in policy were developed more or less consecutively, there is no complete replacement of one generation by another. As will be explained below, there is still a need also for a “Generation 1” command and control policy.

Contaminated site regeneration approaches are nowadays very diverse, but it is useful to make a distinction between three different types of regeneration projects in relation to their economic status and possibilities in financing their remediation and redevelopment. This can be illustrated by the so-called A-B-C Model developed by CLARINET and CABERNET as described in more detail in this book (see Chapter 25 by Nathanail, this book). Depending on the cost of regeneration and the value of the land, sites can be classified according to the A-B-C Model as:

- “Self-developing sites” where the value of the site after reclamation is high compared to the value of the site before redevelopment. Financing of remediation and redevelopment is ensured by private investors (A-sites).
- “Potentially developing sites” which can be redeveloped with certain incentives through public-private partnerships (B-sites).
- “Non-developing sites” or “hard core sites” where development cost too much to attract private investors (C-sites).

It seems that three different styles of policy are required. Self-developing A-sites need a policy aiming at facilitation, potential developing B-sites need a policy aiming at stimulation and the non-developing C-sites need a policy aiming at regulation.

It is obvious that the ultimate aim of a contaminated sites policy is to increase the amount of self-developing sites. The third generation policy is trying to stimulate this. The second generation, where opportunities for public-private financing were created, made the development of potentially developing sites possible. Even with all policy efforts taking away barriers and stimulating sharing of costs and risks by public and private parties, some real hard-core contaminated sites remain unaddressed. To revitalize these or even to manage these, command and control interventions by public authorities are needed. Hence, there is still a niche for the first generation policy approach.

It should be noted that the borderlines between self-developing, potentially developing and non-developing sites can be quite different in countries due to differences in strength of their economy and value of the site. In countries or regions where land is scarce and, hence, where every site has a huge economic potential, the amount of non-developing sites may be small or even non-existent. In other countries self-development of sites may be impossible for economic reasons. These differences are one of the main reasons that countries need to have different policies for contaminated sites. However, irrespective of whether a site is self-developing or not, public authorities should control sites in such a way that the human and environmental risks related to the current use of the site are reduced to an acceptable level.

Generation 2 and 3 approaches to contaminated site redevelopment may involve many stakeholders that can contribute financially and may help to achieve goals set in other areas of policy. There is a substantial body of EU regulations that can be used for dealing with contaminated sites, as shown in Fig. 23.6. If all these could be made to work in concert, remediation of contaminated sites would receive A strong push.

Although national contaminated site policies can be quite different, there is a common need for development of better technologies for detection of contamination and remediation of contaminated soil and groundwater. Current solutions for contaminated site problems become more and more embedded in other socio-economic activities (or problem solutions) at the local and regional scale. Policies promoting decentralised decision-making, minimizing top-down command and control regulation by using a managerial approach, do require that local authorities or private parties that have to manage and restore contaminated sites, have access to adequate tools for assessment, decision-making and to remediation technologies that fit their purposes. Exchange of information and practical experiences in a "learning by experience" process is of utmost importance for the success of decentralised decision-making. So the need for demonstrations of technologies will increase, but in addition there is an increasing need for demonstration of complete remediation and redevelopment concepts.

The developments described above lead to the following conclusions as described in the European Brownfield Agenda (www.eubra.eu):

- There is a need to develop and demonstrate cost-effective site characterisation methods. This may include simple Risk Assessment tools and improved derivation of soil quality standards for decision-making. The HERACLES project may

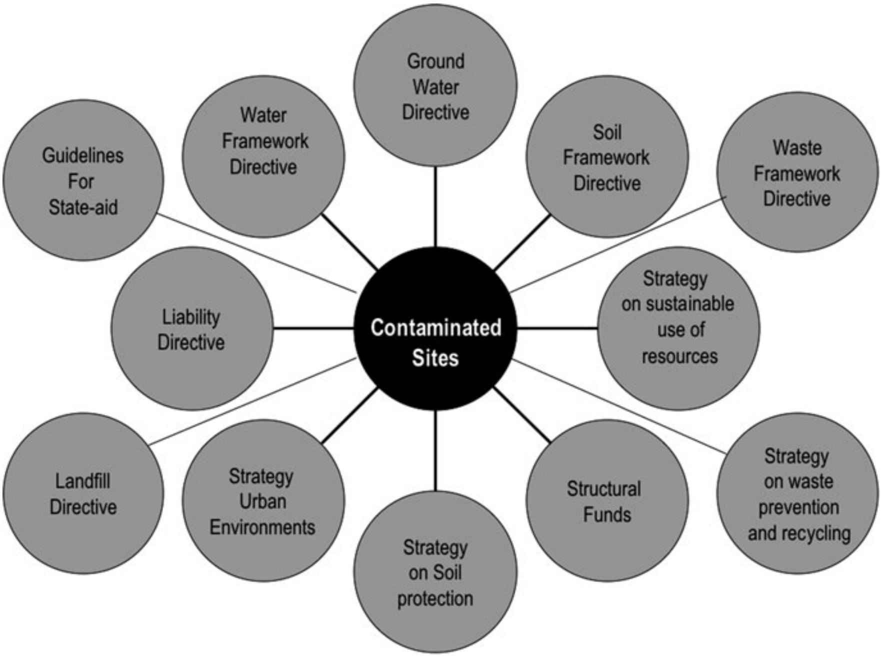


Fig. 23.6 EU policy instruments that can be used for dealing with contaminated sites

be a good starting point for improving the scientific basis for this method (Carlon 2007). Integration of investigation methods and the derivation of trigger values are necessary. Simple, fast and low-cost investigations may require more conservative soil quality standards than comprehensive surveys of a contaminated site.

- There is a need for demonstration of "sustainable site remediation", which are remediation approaches that fit in the social, economical and environmental context of the site.
- There is a need to conserve the collective memory of the experiences from large development projects funded by EU funds like INTERREG. The natural place for a joint collection of experiences in this field in the EU is of course something to be created at the EU level, and this will need some initiative by the European Commission. In addition, national documentation of practical experiences could be made available internationally. It is important that these databases with demonstrations and practical experiences remain "value free" and will not work as a system of technology approval by national or EU authorities.

References

CARACAS – Concerted Action on Risk Assessment of Contaminated Sites in the European Union (1998) http://www.commonforum.eu/publications_caracas.asp. Accessed 24 August 2010

Carlon C (ed) (2007) Derivation methods of soil screening values in Europe. A review and evaluation of national procedures towards harmonisation. European commission, Joint Research Centre, Ispra, EUR 22805-EN

CLARINET – Contaminated Land Rehabilitation Network for Environmental Technologies (2002) http://www.commonforum.eu/publications_clarinet.asp. Accessed 24 August 2010

Common Forum on Contaminated Land in the European Union (2008) http://www.commonforum.eu/. Accessed 24 August 2010

Council of Europe (1972) European Soil Charter, Ref: B(72) 63, Strasbourg 1972

Doves SR, Handmer JW (1995) Ignorance, the precautionary principle and sustainability. Ambio 24(2):92–97

EUBRA (2007) European Brownfield revitalisation agenda. http://www.et-ertel.de/fileadmin/medien/download/EUBRA_Agenda.pdf. Accessed 24 August 2010

EU Groundwater Directive (2006) rective 2006/118/ec of the European parliament and of the council of 12 December 2006 on the protection of groundwater against pollution and deterioration Official J L 372/19, 27/12/2006

EUROSTAT (2003) 2003 results of the LUCAS survey and comparison with 2001 results. Eurostat document ASA/LAND/75, prepared for meeting of the Working group "Land use statistics" of the Agricultural Statistics Committee. December 2003

Ferguson C (1999) Assessing risks from contaminated sites: policy and practice in 16 European countries; Land Contamin Reclam 7(2):33, EPP Publications

Ferguson C, Kasamas H (eds) (1999) Risk assessment for contaminated sites in Europe; Vol 2, Policy frameworks. LQM, Nottingham, ISBN 0 953 3090 1 0

Ferguson CC, Darmendrail D, Freier K, Jensen BK, Jensen J, Kasamas H, Urzelai A, Vegter J (1998) Risk assessment for contaminated sites in Europe. vol 1. Scientific Basis. LQM, Nottingham, ISBN 0953309002

Fowler R (2007) Site contamination law and policy in Europe, North America and Australia – trends and challenges. Paper presented at the 8th meeting of the international committee on contaminated land, Stockholm, September 2007

ICCL – International Committee on Contaminated Land (2007) http://www.iccl.ch. Accessed 24 August 2010

Millennium Ecosystem Assessment (2005). Ecosystems and human well-being: synthesis. Island, Washington, DC

NAS (1983) Risk assessment in the federal government: managing the process. US National Academy of Sciences

NATO Committee on the Challenges of Modern Society (2002) NATO-CCMS pilot study on the evaluation of demonstrated and emerging technologies for treatment and clean up of contaminated land and groundwater. Phase I–III. 1985–2002. http://www.epa.gov/tio, http://www.clu-in.org. Accessed 24 August 2010

NICOLE – Network for Industrially Contaminated Land in Europe http://www.nicole.org/. Accessed 24 August 2010

O'Brien MH (1994) The scientific imperative to move society beyond "just not quite fatal". Environ Prof 16:356–365

Salomons W, Stigliani B (eds) (1995) Biogeodynamics of pollutants in soils and sediments. Springer, Berlin

Shrader Frechette K (1996) Methodological rules for four classes of scientific uncertainty. In: Lemmons J (ed) Scientific uncertainty and environmental problem solving. Blackwell science, USA. ISBN 0-86542-476-4

Siegrist RL (1990) Development and implementation of soil quality and clean-up criteria for contaminated sites. In: Arendt F, Hinselveld M, Van den Brink WJ (eds) Contaminated Soil'90; third international KfK/TNO conference on Contaminated Soil; Kluwer Academic, Dordrecht/Boston/London

Soil Communication COM(2002) 179 final. Towards a thematic strategy for soil protection. http://ec.europa.eu/environment/soil/index_en.htm. Accessed 24 August 2010

Soil Communication COM (2006) 231. http://ec.europa.eu/environment/soil/index_en.htm. Accessed 24 August 2010

UK DETR Circular 2 (2000) Contaminated land: implementation of part IIA of the environmental protection act 1990. Department of the Environment, Transport and the Regions. Published 20

Van-Camp L, Bujarrabal B, Gentile A-R, Jones RJA, Montanarella L, Olazabal C, Selvaradjou S-K (2004). Reports of the technical working groups established under the thematic strategy for soil protection. EUR 21319 EN/4, 872 pp. Office for Official Publications of the European Communities, Luxembourg

Vegter JJ, Lowe J, Kasamas H (2002) Sustainable management of contaminated land: an overview. Austrian Federal Environment Agency, on behalf of CLARINET. ISBN3-85457-674-9

Visser WJF (1993) Contaminated land policies in some industrialised countries. TCB R02, The Hague

Vlek CAJ (1990) Beslissen over risicoacceptatie (in Dutch). Gezondheidsraad A 90/10 Den Haag

Vlek CAJ (1995) Risk assessment, risk acceptance and risk management: a psychological decision theorist's view. In: Van den Brink WJ et al (eds) Contaminated soil '95. Kluwer Academic, Dordrecht, pp 565–579

Water Framework Directive. Directive (2000) 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for Community action in the field of water policy

Weinberg AM (1972) Science and TRANS-science. Minerva X 2:202–222

Wynne B (1972) Uncertainty and environmental learning: reconceiving science and policy in the preventive paradigm. Global Environ Change 2:111–127

Chapter 24
A Stakeholder's Perspective on Contaminated Land Management

Lida Schelwald-Van der Kley, Johan de Fraye, Robert Pentel, Lucia Buve, Roger Jacquet, Hans L.A. Slenders, Ian Heasman, Steve Wallace, and Markus Ackermann

Abstract NICOLE, the Network for the management of Industrially Contaminated Land in Europe is a leading forum on this matter in Europe, promoting co-operation between industry, academia and service providers. NICOLE's goal is to enable European industry to identify, assess and manage industrially contaminated land efficiently, cost effectively and within a framework of sustainability. This chapter provides an introduction to NICOLE's philosophy and addresses a number of topics that are relevant to NICOLE's industrial members throughout the EU. Topics range from consideration about consistency in legislation to innovative, risk-based and sustainable approaches for the management of operational sites, megasites and Brownfields. The chapter outlines the existing concerns and constraints that may hinder cost-effective contaminated land management, proposes solutions and provides examples of promising and proven methods or best practice. NICOLE's industrial members have developed and put into practice a range of techniques and methods to reduce, alleviate and prevent contamination of soil, surface water and groundwater. A greater knowledge and understanding of the forms and nature of contamination has allowed the refinement and improvement of remediation approaches used. NICOLE and its members advocate the view that remediation should be sustainable, i.e. that there be an acceptable balance between the effects of undertaking remediation activities and the benefits the same activities will deliver, in terms of environmental, economic and social indicators.

L. Schelwald-Van der Kley (✉)
Envision-S b.v., Lelystad, The Netherlands
e-mail: mail@schelwald.nl

The opinions expressed in this chapter do not necessarily reflect the view of all individual members or organisations within NICOLE, but are generally supported by most of NICOLE's industrial members.

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_24,

Contents

24.1 What is NICOLE and What is This Chapter About?

NICOLE, the Network for the management of Industrially Contaminated Land in Europe began its activities in 1996, as an EU project, funded under the Framework 4 programme, aiming at a better linkage between research and industrial needs. In 1998 NICOLE became a self funding network focussing on the promotion of *Risk-Based Land Management* (RBLM) approaches.

NICOLE comprises two subgroups: Industry and Service Providers. The Industry Subgroup currently counts 30 members. In addition NICOLE has individual members from organisations such as universities and governmental organisations. A common trait of members of the Industry Subgroup is that they own the problem (the responsibility for contaminated land). The Service Providers' subgroup represents both internationally operating consulting firms and those that have a regional coverage. NICOLE maintains an active programme including two full conferences per year and many individual working group and sectoral group meetings. The main benefits of membership are a cross-fertilisation of ideas among sectors and countries and transfer of knowledge. Particularly important is early knowledge of European initiatives and the opportunity to take part in an international forum which can provide a concerted technical influence at a European level. To quote one of NICOLE's service provider members: "NICOLE is a group of well qualified and highly skilled people, where competition between consultants and self interest is put aside and we all work for the same targets. This is an excellent platform for sharing knowledge and know-how, and improving practice inside and outside NICOLE. When a Nicole member takes part in a project you can expect that all

Table 24.1 Selected NICOLE Publications from 2005 Onwards

2008	*Report of the NICOLE Workshop:* Environmental Decision Support Systems 9-10 October 2008 Madrid, Spain. See http://www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 17(2):275–314
2008	NICOLE Position Paper (2008) A note on proposals for increased soil and groundwater monitoring under the proposed new IPPC Directive http://www.nicole.org/documents/DocumentList.aspx?l=9andw=n
2008	Report of the NICOLE / SAGTA Workshop: Sustainable Remediation 3rd March 2008, London, UK. http://www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 16(4):381–403
2007	*Report of the NICOLE Workshop:* Using baselines in liability management – what do upcoming Directives require from us? Brussels, Belgium. See www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 16(3):277–306
2007	*Report of the NICOLE Workshop*: Redevelopment of sites – the industrial perspective, Akersloot, the Netherlands. See www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 16(1):50–75
2007	NICOLE Position Paper (2007) Comments on the proposal for a directive of the European parliament and of the council on waste. NICOLE, TNO, The Netherlands. http://www.nicole.org/documents/DocumentList.aspx?l=9andw=n
2007	NICOLE Position Paper (2007) Concerning European commission communication "Thematic Strategy for Soil Protection" COM(2006)231 final ("strategy") and Proposal for Directive of the European Parliament and of the Council establishing a framework for the protection of soil and amending Directive 2004/35/EC ("directive"). Available from NICOLE Secretariat, TNO, the Netherlands. http://www.nicole.org/documents/DocumentList.aspx?l=9andw=n
2006	*Report of the NICOLE 1996-2006 Ten Year Network Anniversary Workshop*: Making Management of Contaminated Land an Obsolete Business – Challenges for the Future. 5 to October 2006, Leuven, Belgium. See http://www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 15(2):261–287
2006	*Report of the NICOLE Workshop*: Data Acquisition for a Good Conceptual Site Model 10–12 May 2006, Carcassonne, France. See www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 15(1):94–144
2006	*Report of the NICOLE Workshop*: The Impact of EU Directives on the management of contaminated land, 1–2 December 2005, Cagliari, Sardinia, Italy. See www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 14(4):855–887
2005	*NICOLE Report*: Monitored Natural Attenuation: Demonstration and Review of the Applicability of MNA at Eight field sites, http://www.nicole.org/publications/library.asp?listing=5 (summary). The full report can be ordered from the NICOLE Secretariat
2005	*NICOLE Report*: The Interaction between Soil and Waste Legislation in Ten European Union Countries sites, http://www.nicole.org/publications/library.asp?listing=7 (summary). The full report can be ordered from the NICOLE Secretariat

Table 24.1 (continued)

2005	*Report of the NICOLE Workshop*: State of the art of (Ecological) Risk Assessment, 15-16-17 June 2005, Stockholm, Sweden see http://www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 14(3):745–773
2005	*Report of the NICOLE Workshop*: Unlocking the Barriers to the Recovery of Soil and the Rehabilitation of Contaminated Land. 15–16 November 2004, Sofia, Bulgaria see www.nicole.org/publications/library.asp?listing=1; Land Contamin Reclam 14(1):137–164

available knowledge is incorporated in the project. In addition, the co-operation between service-provider and industry groups gives me a better understanding of the problems, needs and wishes of (multinational) industry. NICOLE also acts as my antennae for new developments (legislation and innovation) in the field of contaminated land management."

The exchange of views, knowledge and best practices with other stakeholders (often regulators) takes place at NICOLE workshops, which are held twice a year. Other forms of interaction and dissemination include NICOLE applied research projects and Working groups (often related to Member States' implementation of one of the European Directives), as well as interaction with other European networks such as the Common Forum on Contaminated Land and supporting participation in research programmes, such as EC Framework Programmes, and their review panels. NICOLE's work over the past 15 years has led to a wide range of freely downloadable publications and reports; some examples are listed in Table 24.1.

24.2 The Road to Sustainable Risk Based Land Management

To illustrate how NICOLE's industrial problem holders view the contaminated sites "landscape" that they operate in, as well as the progress that has been made in the sector during the existence of NICOLE, Appendix contains two "Road maps to Sustainable Risk Based Land Management". The first one was originally developed in 1998 as part of a NICOLE initiative to help those concerned with developing contaminated land management practice to visualise the issues they were facing. The map depicted contaminated land as terrain that was relatively unknown, with alternate routes and possible perils through the landscape to attain the desired goal of successful restoration. It was envisioned that the most successful path was one travelled together using Risk Assessment and Risk Management to negotiate the terrain and reach the goal. The second roadmap is an update of the 1998 one, made in 2006. A comparison of the two clearly illustrates the changes that have taken place within this time frame, reflecting experience and new knowledge. The people who need to use the map remain the same, with assistance still provided by the service and academic communities. However, the obstacles have reduced in size as experience has provided increased confidence in the technological approaches to contaminated land management and the level of doubt about approaches and science used has been reduced. At the same time, this experience has shown that some routes to achieving fit-for-use land have been found to be less desirable than they were believed to be in 1998. These roadmaps illustrate that contaminated land management as a sector has developed significantly in the last ten years as a result of greater knowledge and experience as well as the benefit gained from different stakeholders working together across Europe, a process that NICOLE has helped to initiate and develop. One end result of such cooperation has been the realisation that effective management of contaminated land requires more than just a technical solution in order to be successful, but also strategic planning.

24.3 A Strategic Approach to Contaminated Site Management: The End State Vision

When faced with a contamination issue such as a groundwater plume, many remediation project teams concentrate on the immediate issue at hand and a technical approach which offers an actual solution or improvement. However, also the strategic discussion on long-term use of the site needs to be included in the choice of the remediation.

A strategic approach requires the definition of an endpoint for the remediation of the site, which in turn requires an answer to the question "what does the future of the site look like?" For example, the plan could be to come to full site closure with no remaining liabilities, or, alternatively, to continued operation with no risks for site operators and downstream users. An "End State Vision" relates to a very high level concept. Consequently, its development requires the gathering of many stakeholders' views and a consideration of issues beyond pure technical solutions.

This end state vision requires that the remediation team has knowledge of the future of the site and of the wishes and needs of the stakeholders. Especially for large (former) industrial sites this is a complex exercise, which will lead to an "optimised" end vision as it is generally impossible to satisfy all stakeholders at all fronts. Once the End State Vision has been formalized, the remediation strategy and all techniques and methods associated with it can be developed. In many cases the remediation team has a range of techniques to select from. The choice of remedial technique depends on many factors which in part are non-technical. The selection is influenced by the experience of the project team, the consultant and local regulator. Also the timeframe and "predictability" of the technique influence the decision. Too often a quick solution such as dig and haul, is preferred over a longer term, more sustainable, solution because the stakeholders are uncomfortable with the -perceived- uncertainty. For many sites the technical approach is just a minor part of the solution. There are many other non-technical considerations that can have a major impact on a project: examples include financial considerations (e.g., the value of property, the cost of remediation), social considerations (e.g., the interests of neighbours, agencies, employees, shareholders), temporal considerations (e.g. time of occupancy of a site, imminent sale, redevelopment schedule), and legal considerations (e.g., existing or upcoming legal schemes or contractual arrangements between parties). Technical possibilities (geology, chemistry, biology, technology) also play a role, but are mostly subordinate to the other aspects of a contaminated site project.

Part of the strategic discussion is the long-term management of the site. In case there is residual contamination, reassurance should be given on who is going to manage and monitor this and make sure that the contaminant situation is and stays in line with the future use of the site. In some countries land spatial planning tools and land-use restrictions can be filed with authorities, reassuring that future users cannot develop the site for a more sensitive use than the residual contamination would allow.

For some sites the ownership changes during this route to the "End State Vision". Typically, when the ownership of contaminated or remediated land changes, there is a transfer or sharing of liability. This may be explicit by contractual agreement between the vendor and the purchaser, or it may be implicit based on the context of the land transaction and local regulations. Many large land owners have been in existence for some time and are also seen as having "deep pockets". For this reason they may be cautious about transfer of ownership to organisations that they see as less (financially) stable, or likely to take actions such as land use change which is outside their control. This is because they fear that any consequent liability will still remain with them as original polluters. This can create a tension within organisations who on the one hand wish to securely manage their own liability in the long term, but on the other hand would also like to see a return on these assets and their productive reuse. In sales contracts these long-term contractual obligations between parties are detailed to protect both parties; occasionally a "dowry" or bank guarantee is paid for the transfer of a site to a redeveloper which is released to the new owner in tranches according to the stepwise clean-up and reuse of the site.

In short, a well managed, successful project closure requires a holistic approach to remediation, which provides the technology required to meet all or most of the stakeholders' needs. Very often this balance must acknowledge that stakeholder needs for a particular site can be addressed without the technological goal of complete remediation having been achieved. This may mean in some cases that the level of risk that is acceptable for a given purpose and that can be achieved within the constraints of time, budget and technology has to be defined. This process is illustrated by the case study of the management of manufactured-gas plants In France, described below.

Manufactured gas plants (MGPs) were once widespread across Europe. Based on the distillation of coal, this process produced a purified gas, mostly methane, which was then stored in gasometers before being injected into the gas network. Such gas plants were in operation in France from 1798 until the late 1960s, when the development of natural gas reserves resulted in the gradual closure of all MGPs across the country, the last of which ceased operations in 1971. However, although the surface installations were dismantled during the decommissioning of the sites, by-products from the distillation of coal and gas purification processes were not always eliminated. These by-products may, for example, have been confined in tanks or absorbed by soil. The need to manage old MGP sites in a coordinated, coherent way was highlighted in France during the 1990s at the time the issue of contaminated sites emerged in France, following the discovery of tar tanks containing by-products from the production process of manufactured gas on a site owned by the utility company *Gaz de France (now GDF SUEZ).*

In order to coordinate action throughout the national territory, *Gaz de France / GDF SUEZ* identified and prioritized the 467 sites of old gas plants for which it was responsible and developed a Memorandum of Understanding (MOU) in conjunction with the Ministry of Ecology and Sustainable Development of France. This MOU pertains to the control and monitoring of the rehabilitation of these obsolete sites and was signed on April 25, 1996, for a period of 10 years. The methodology laid down within the MOU protocol was used to rank sites according to their risks for the environment, on the basis of the use of the site, the vulnerability of groundwater and surface water to contamination, the presence and type of contamination on the site, et cetera. In 2001, *Gaz de France / GDF SUEZ* also conducted a "semi-generic" study for rehabilitation of the MGP sites throughout France. It defined remediation objectives for which the level of human and environmental risk would be acceptable given the intended end use of the site.

As another aspect of its policy on former industrial sites, *Gaz de France / GDF SUEZ* systematically carries out initial site investigations prior to any sales transaction, transfer or redevelopment of these sites. This is in fact the principle activity of the former MGP sites management policy of *Gaz de France / GDF SUEZ.* A lot of these sites are to be sold, and the remediation objectives depend on the final project and land use needs of the buyer. The cost of the remediation required is negotiated between the two parties. The management of the monitoring and the remediation of other sites remaining under the ownership of *GDF Suez* also depend on their ongoing use.

Some of the key results of the implementation of the MOU protocol are as follows:

- all 467 sites have been investigated;
- 332 hectares of land have been remediated and reused in urban planning projects;
- investigations and works led to the draining and filling of tanks from 270 sites and the safe management of over 715,000 tons of waste;
- surveys and regular monitoring of groundwater have been carried out at 268 sites: 93 sites are still subject to such actions.

The end of the MOU period was reached in 2006 and an output document was co-signed by the Minister of Ecology and Sustainable Development of France and the Chairman of *Gaz de France*. For the remaining 93 sites under investigation, a coordinated decision on future activities (i.e. an extension, reduction or suspension of the current water monitoring programme) remains to be made.

24.4 Improving the Efficiency of Site Assessment

Since its beginning in 1996, NICOLE has advocated and provided support for a Risk Assessment-based approach as the fundamental basis for sound and sustainable management of contaminated sites. This is because it allows for a flexible approach to decision making whereby corrective action is tailored to site-specific conditions. This leads to cost-effective solutions that are appropriate to treat the actual risks posed by contamination.

Today Risk Based Land Management (RBLM) is considered the best available strategy for dealing with the problems posed by contaminated sites by most EU Member States (Carlon and Swartjes 2007) and many other developed countries in the world. Consequently, many countries have adopted the risk-based principle in their environmental policy on contaminated land, although the harmonisation of both human and ecological Risk Assessment methodologies across Europe is still an issue at hand (Carlon and Swartjes 2007).

Efficient site characterisation and management of the data associated with it are recurring topics within NICOLE. There is a balance between necessary site characterisation costs to provide sufficient knowledge for risk based decision making, and excessive site characterisation that is carried out purely for compliance purposes without any real improvement in decision-making. This excessive cost produces information of low value, and is a misuse of resources which could be better directed elsewhere in site management. Figure 24.1 illustrates this idea that there is an optimum level of site characterisation effort: too little and overall site management costs are driven higher – for example because incorrect decisions are made; too much and the additional value of information provided does not improve the site management and simply adds cost to the project. Figure 24.1 also introduces the idea that this cost curve is affected by other factors. A decision making process that is perception-driven as opposed to being knowledge-driven will tend to lead to a more expensive

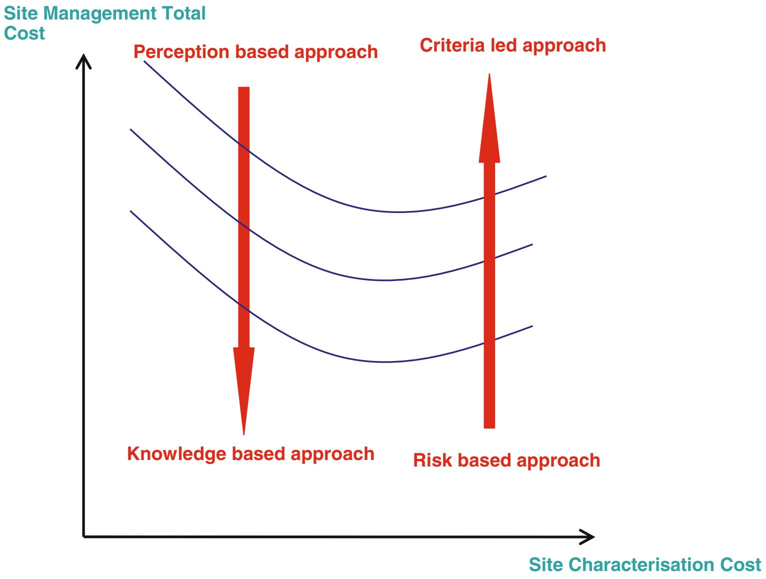


Fig. 24.1 Site characterisation cost curves – conceptual diagram

project; site characterisation that is driven by compliance with particular sets of prescribed criteria will be less effective, and so more expensive, than one based on effective Risk Management decision making. Technical innovation may also reduce costs, in particular iterative approaches to site characterisation that allow "real-time" decision making, such as the US EPA Triad approach (www.triadcentral.org).

It must be noted that no single solution can provide the optimum management approach for all contaminated sites. The approaches to investigation typically evolve alongside the refinement of the respective site conceptual model, which serves as the ultimate basis for the subsequent Risk Assessment and remediation strategy. It is important to ensure that such a site conceptual model does indeed represent all of the relevant source-pathway-receptor linkages.

24.5 Remediation of Contaminated Sites and Waste Management

The definition of waste hinges on the term "discard" and "intent to discard", although "discard" itself is not defined in the original 1975 Waste Framework Directive. Over the last 25 years a large body of case law has arisen to help to better define waste. Whilst soils and other materials from Brownfield reclamation and contaminated land management have long been considered waste when sent to landfill, it has only been in about the last 10–15 years that these materials are sometimes considered waste when managed and reused on site. A consensus has emerged that significantly contaminated soils normally become waste at the point of excavation, but there has been much confusion and variability in waste regulation practice over other materials such as marginally contaminated soils, made ground, natural ground and other materials.

Whilst being waste may simply appear to be a label, it in fact carries significant legal responsibilities. Materials with the waste label, even though they may be safe and suitable for use, may be subject to stigma and cause blight at the site where they are reused. Once such materials have been classified as waste there is also much inconsistency over the point at which they cease to be waste. This variably may be at different points in the value chain either on site, off site (at soil treatment centres), at receiving sites, or when the material reaches its final resting place in the ground.

Some countries have attempted to dually regulate soil as waste (where definitions are hazard based) alongside risk-based soil management processes. This typically leads to complexity and contradiction and to non-optimal resource management. In some cases this has led Member States to put in complex arrangements in order to reconcile waste law with planning law and national law in relation to contaminated land. Others have simply declared soil as a resource and never treat soil as waste except where it is clearly discarded to landfill. In between these extremes there is a whole spectrum of approaches.

Overall the complexities and variability around soil as waste can have significant impacts on contaminated land and Brownfield project management, costs and the ability to reuse materials. Opportunities to reuse treated and untreated soils are decreased and the burden on landfill and the use of primary, quarried, replacement intensified.

The revised Waste Framework Directive will be transposed by December 2010. The new Directive's focus on recycling and reuse means that all materials including soils may be more clearly defined as by-product rather than waste, and there is more certainty about when materials cease to be a waste. It has also sought to address some specific issues around soil as waste. For example it counters the Van Der Walle ruling excluding unexcavated contaminated soil from waste control. It also allows excavated uncontaminated soil from construction processes to be reused on the site of origin outside waste control. NICOLE is hopeful that appropriate and consistent transposition of the new Directive will improve the current complexity and inconsistency that prevails.

24.6 The Power of Natural Processes

Experiences with subsurface remediation worldwide have made clear that our society is not able to achieve a complete removal of contaminants in the subsurface within time frames of a few decades. Moreover, a complete removal is in most cases technically impossible and the costs of full removal are often not in line with the benefits and risks involved. In addition the complete removal of contamination could be contrary to the general idea of a smaller footprint in all human activities (c.f. remediation in a sustainable context. Therefore remediation should not be considered as an end goal in its own right, but rather as a necessary effort in maintaining the useful properties of land – a non-expandable good. Fortunately, it is possible to make use of the naturally occurring processes in the subsoil. *Natural Attenuation* (NA) is increasingly recognized as a process that can result in gradual loss of contamination, especially in plume areas of contaminated sites.

In 2000, NICOLE set up a data sharing program for industrial sites to investigate the possibilities for *Monitored Natural Attenuation* (MNA). Eight industrial partners and their consultants carried out investigations at their sites as part of a NICOLE MNA demonstration project. The results were reviewed by 12 independent reviewers from universities and authorities. The general outcome was that MNA is applicable and effective at many sites. This demonstration project underlined the fact that MNA is a cost-effective option that can be used to obtain satisfactory results. MNA is often not a standalone option, although it could be. In many approaches it is part of a risk-based site management plan. It is the view of NICOLE that MNA should be incorporated into the various soil and groundwater policies as one of the tools to manage historical soil and groundwater contamination and to support the objectives. Generally speaking it is necessary that natural degradation processes are taken into account in Risk Management solutions. (Monitored) Natural Attenuation is described in detail in Peter et al. (Chapter 22 of this book).

Some consider MNA as a panacea, a magic solution to contamination problems. Detractors, however, will argue that this is only the case over significant timescales. Nevertheless, Natural Attenuation processes reduce the impact of past, present and future soil and groundwater contamination. Contrary to perceived wisdom, in some instances Natural Attenuation can be fast, for example in the case of a limited impact of readily biodegradable contaminants. In a specific recorded instance of a new spill, management of the contamination through MNA occurred faster than the administrative timeframe for the management of the contamination.

Obviously, when MNA eliminates the contamination to such extent that the risk to humans and the environment are acceptable there is no reason for further action. For instance, if contamination is limited to an operating industrial site and the risk of exposure of the workers is acceptable, it is not more efficient to invest in prevention measures then in remediation. Remediation would not improve the safety of workers, in fact it could even supply a (temporary) exposure pathway. On the other hand, investment in prevention measures will reduce the risk of occurrence of new incidents and their impact on the subsoil, improving the workers' safety and increasing the level of environmental protection, while MNA will address any historical contamination.

24.7 Managing Megasites; An Integrated Approach

In broad terms, megasites can be described as large conurbations of sites where contamination has arisen independently, but in which contamination is so widespread that the problem extends across the whole area rather than individual sites. Such sites can cover large areas, but they could also be smaller areas where several sites with different operators and owners are linked by overlapping and/or shared land and water contamination problems. As contamination develops and spreads, the environmental problems caused by these individual sites become inter-linked. The management of such megasites may therefore be facilitated by taking an overarching approach, rather than trying to deal with each site on an individual basis. This is especially true for groundwater, as contaminants are not limited to individual site boundaries. But even where contamination problems in the subsoil remain relatively discrete there may be economic, social or environmental advantages in taking an integrated approach.

The management of megasites poses major challenges to the current paradigm of contaminated land management, for example by advocating a multi-site owner response. . Such an approach should be supported by soil and (ground)water policy and legislation, with a recognition of risk-based decision making specific to megasites. The case study of the *Port of Rotterdam industrial area* provides a good example of an integrated megasite approach, where a large number of interests carry out operations on a number of sites. The area around the Port of Rotterdam covers approximately 3,000 industrial sites. Thousands of contaminant sources are thought to be present. The entire port area, the municipality of Rotterdam and the polder areas is being considered as a megasite. Further examples are available in a NICOLE Workshop report on megasites (NICOLE 2004).

The development of the Port of Rotterdam started in the 17th century and had its main period of expansion during the 19th and 20th centuries. Currently, the industrial sites in the port and surrounding area cover an area of approximately 10,500 ha, managed by the Port of Rotterdam Authority. The area itself is in a dynamic phase. Old port basins are being redeveloped to accommodate the urban expansion of the city of Rotterdam and new port areas are being constructed on the seafront. Among the main activities that take place in the Port are the trans-shipment and processing of bulk goods such as oil, chemicals, coals and ores. As a result of these industrial activities both soil and groundwater have become contaminated with a wide range of contaminants, some of which reached the groundwater and, depending on the specific geohydrological situation, might form extensive plumes.

Since the late 1980s the Port of Rotterdam policy has been to investigate and record initial soil quality when issuing a lease contract to companies. The basic principle is that when a company leaves the site and the lease contract is ended, the Port of Rotterdam will accept the site back only in its original state. In practice a functional, fit for purpose, remediation is allowed. As financial security for the allowance of residual contamination and to compensate for capital loss, an integral claim is maintained. If the company is sold to another company, the selling company will be held liable for any remediation costs, although the selling and buying companies may sign a mutual agreement whereby the buyer takes over the liability.

Although some companies are eager to have their sites quickly remediated, the Port of Rotterdam views the contamination issues in a greater context. Contaminants seldom respect site boundaries, certainly not when the deeper aquifers are contaminated. As a good landlord, the Port of Rotterdam advocates a more coherent solution, in which soil, groundwater and surface water are viewed as an integrated system. Remediation of the subsoil remains the responsibility of the individual companies and the contaminated groundwater is dealt with in a regional or clustered approach. Currently the Port of Rotterdam is working towards an agreement between both public and private partners to participate and share the costs for this. The Port's ultimate goal is to manage the soil and groundwater quality in the port area in the most optimal manner, thereby ensuring its economic activity in the coming years.

From the point of view of the Groundwater Directive (GWD), the contaminated groundwater beneath the Port of Rotterdam can be seen as one plume resulting from contaminated sites and point sources. The key goals are to keep the plume from spreading further and not to present a risk to human health and the environment. National and local legislation has been designed to limit the risks of individual point sources and also gives the Port of Rotterdam space and time to use Natural Attenuation for the risk-based management of groundwater plumes. The surface water, the Pleistocene deep aquifer and the downstream polder areas are all potential

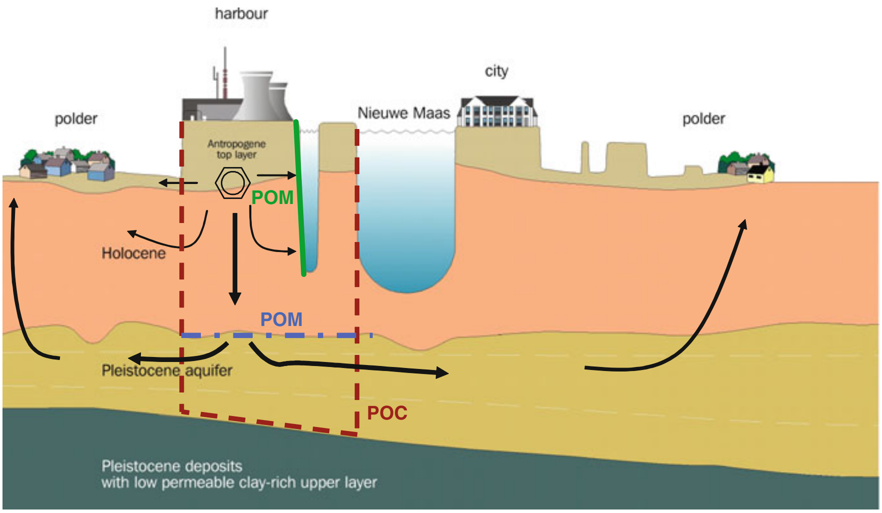


Fig. 24.2 Conceptual model for the Rotterdam Port area, indicating the groundwater flow (*black arrows*), the planes of monitoring (POM) and the plane of compliance (POC)

receptors in danger of receiving contamination at the Rotterdam megasite due to the direction of groundwater flow (Fig. 24.2).

The overall quality of the surface water has been found to be only "marginally" affected by the contaminant plumes in groundwater. Therefore the focus is on groundwater as the most important receptor to protect. In accordance with the requirements of the draft GWD, the *DCMR Environmental Protection Agency*, as the environmental authority of the Port of Rotterdam area, and representatives of the industry in the Port of Rotterdam (united in the *Deltalinqs organisation*) have together developed an approach to verify that groundwater plumes from contaminated sites do not expand.

To protect the surrounding groundwater body, *planes of monitoring* (POM) and *planes of compliance* (POC) have been defined. The POC forms the boundary between the partly impacted groundwater zone beneath the port and the surrounding groundwater zone that needs to be protected. The lower boundary of the POC is defined by the boundary between the Pleistocene aquifer and the low permeable clay-rich layer beneath (aquitard). The groundwater zone enclosed by the POC provides time and space for Natural Attenuation processes and the management of the plume. To protect the surrounding groundwater zone, contaminant concentrations are compared to legislative standards at the POC itself, and if necessary followed by adequate remediation measures. The two POMs are intended to monitor the trends and the processes in the plume and function as an early warning for a possible future impact to the POC. In case of an uncontrolled plume expansion, measures can be taken at an early stage. This approach is currently being tested at pilot scale with a number of companies in the port area.

24.8 Brownfields: A Blessing in Disguise?

CABERNET defines Brownfields as sites that:

- have been affected by the former uses of the site and surrounding land;
- are derelict and underused;
- may have real or perceived contamination problems;
- are mainly in developed urban areas;
- and require intervention to bring them back to beneficial use.

As a network of industrial problem holders, NICOLE's principal interest in Brownfields is in relation to the closure of industrial sites which will then meet the CABERNET definition.

Europe and the industrial world on the whole is still cluttered with plots and sites in mostly suburban areas. Their remediation and redevelopment had so many hurdles in the past that many of them have remained in a state of advanced dilapidation of their buildings as a visible sign of their abandonment, both economically and societal., With city centres spreading out and economy shifting to cleaner industry and commerce, some of the Brownfields are today in a desired location and merit a new use.

Yet, generally speaking, Brownfield sites are still regarded as a problem requiring a solution rather than offering an opportunity. The counterpoint to this view was addressed in the NICOLE workshop *"The Redevelopment of Sites from an Industrial Perspective"*, held in Akersloot, the Netherlands, in June 2007. This workshop highlighted the benefits of the reuse of Brownfield sites for economic, social and environmental regeneration. Examples were given of the advantages to local communities, governments, vendors and developers of Brownfield sites, from site to city scale. As introduced above, Brownfield sites are often situated on premium locations (A-sites) for housing or other high-value development: it was shown that their development releases funds that can be used to deal with the historic legacy of soil and groundwater contamination.

Nevertheless, industrial site owners are very much aware of the issues associated with selling and redeveloping contaminated land in particular for sensitive uses, such as housing. In particular, many industrial companies are sensitive to the possibility that, even with contractual and legal safeguards in place, it could be possible for them to be confronted with the responsibility for a post-divestiture contamination problem. Thus potential future liability is a major concern, not simply from a financial perspective, but also from the point of view of reputation damage, that may well have been hard won. In many cases, financial concerns take a secondary role against the adverse publicity that can surface if major owners are retroactively called back to remediate sites which have been redeveloped. More specific concerns include:

- Institutional controls – How can initial and subsequent uses of property be controlled to prevent uses that are not encompassed within a Risk Assessment framework? In many cases remediation objectives are set for the immediate property use. Future land uses may not be in line with these remediation objectives, while land use change often is beyond the control of the initial property owner. With the question of successors in title this control becomes even more difficult to exert.

- Certification – In all EU countries ultimate liability rests with the polluter (the polluter pays principle) and even in cases where remediation has been accomplished the potential is always there for recurrence of liabilities.
- Financial security – The question which mechanisms are present, or could be developed, to safeguard current owners when sites are inappropriately developed, or successive owners default on obligations and require additional remediation activities.
- Development incentives – Currently the only incentive in many cases is the inherent land value, although remediation costs often mean that the land value of many Brownfields will be negative. The question is which incentives can be introduced to make redevelopment more attractive.

With greater regulatory interest and demand for Brownfield development in some countries, land owners are facing increasing pressure to surrender sites for development. In case of inappropriate development, the negative aspects of developments at the site are frequently attributed to the land owner. Therefore, it is not surprising that many owners do not proactively pursue development of their land. This situation can best be illustrated by an examination of the risks pertaining to the potential negative aspects of incorrect development.

At the beginning of 2008, NICOLE created a Brownfield Working Group, whose main objective is to develop an approach to support the transfer of contaminated sites. The purpose of the group is to stimulate responsible corporate land owners to divest land with confidence and the certainty of a proper transfer of liability. This would be approached through evaluating current transaction processes, in order to identify those areas that may give rise to concern.

24.9 Redeveloping Contaminated Industrial Sites: A UK Developers Perspective

As the birthplace of the industrial revolution, the UK can claim a legacy of contaminated industrial sites second to none. Over the last 25 years, a combination of factors has resulted in a progressive approach to the redevelopment of these sites, perhaps unmatched elsewhere in the world. The UK approach has been driven by greenbelt and urban planning policies, high population density, high housing demand, environmental regulation, tax incentives, public reluctance to new development (as opposed to redevelopment) in their neighbourhood, British pragmatism and even a touch of foresight. After risk based site remediation has been completed to make sites safe, homes, shops, offices and factories are now routinely built on sites that were previously chemical works, gas works, oil depots, and manufacturing sites.

The UK was an early adopter of the "suitable for use" philosophy for site reuse, avoiding the multi-functionality approach. Industrial site remediation in the UK is decentralised and strongly development led. This moves remediation costs away from the public purse and into the development feasibility calculations. Hence, these legacy costs are factored into current land values. The government recently adopted a *National Brownfield Strategy* to deal with dereliction and accelerate renewal. It also has a target of building 60% of all new homes on "previously developed land" (a term commonly but not necessarily accurately used synonymously with the term "Brownfield") that has long been exceeded.

While not all Brownfield sites are contaminated, or even industrial, a substantial proportion is. Redevelopment of industrially contaminated sites is mainly controlled through the "development control" or "land use planning" system. Contaminated sites are a material planning consideration and hence needs to be fully considered as part of any application for redevelopment. Sitting alongside this and also requiring consideration is Part IIA of the Environmental Protection Act 1990, national legislation designed to protect human health and the environment from the worst contaminated sites. A voluntary approach to remediation is encouraged with direct regulation seen as the last resort. The combination of the "carrot" of redevelopment profits and the "stick" of legislative intervention has in general been effective.

Risk Assessment matches well with the UK's pragmatic approach to remediation, and has long been established. The application of *in situ* and *ex situ* chemical, physical, biological and thermal remediation technologies is now commonplace in the preparation of post-industrial sites for redevelopment. The requirements of the Landfill Directive implemented in 2004 gave a substantial boost to on site remediation, off site soil treatment centres (already seen commonly in Northern Europe) as well as our home grown CLUSTER concept.[1] The removal of landfill tax exemption for contaminated sites in November 2008, in combination with the landfill tax escalator is anticipated to give the remediation sector a further boost, although the effects are as yet hard to discern because of the economic downturn. Remediation sites can also get tax benefits via Remediation Corporation Tax Relief, currently equivalent to an additional 14% of the remediation costs for developers. However, problematic waste streams still exist for which there are no available technology solutions or where the available solutions have significant other environmental impacts (e.g. thermal treatment). There is a danger that prohibitively high landfill costs for such waste streams will result in some sites being made safe and "moth-balled" rather than being returned to beneficial use.

Of course the UK approach is not a panacea, and there are some problems in this approach to industrial site remediation that are home-grown. Among these are the UK interpretation of the definition of waste, the legislative requirements of the "waste label" for soils, and identifying when materials cease to be a waste.

---

[1] In CLUSTER contaminated soils from "donor" sites are treated at a "hub" site (where the remediation technology is temporarily based) and clean soils returned to the donor site.

This can result in dual regulation problems of applying waste legislation and risk-based contaminated sites legislation to the same activity. Hopefully, after years of concerted effort we are moving closer towards resolution in this area.

One consequence of having a redevelopment-market-led approach to industrial site remediation is that a downturn in the construction sector, which the UK entered in the last quarter of 2007, could have significant consequences for the UK remediation sector. This, at a time when Environmental Liability transfer is becoming a key driver in the divestment of land, will make remediation of industrially contaminated sites even more challenging in the future.

Procedures for dealing with Brownfields are described in detail in Nathanail (Chapter 25 of this book).

24.10 A Sustainable Future?

24.10.1 Sustainable Approaches

The main challenge for the future is to ensure that management of contaminated sites, like that of any other complex issue faced by our society, fits within a framework of *sustainability*. Ensuring that activities undertaken today will not be detrimental to environmental quality in the future, nor will they restrict activities in the future, is an increasingly important political technical and practical issue.

NICOLE has held three workshops, in 2003, 2008 and 2009 to discuss sustainable approaches to remediation with representatives from industry, service providers, researchers and regulators to explore the opportunities, challenges and barriers posed by sustainable site management. It was clear in 2003 that the meanings ascribed to terms such as “sustainable” or “sustainable development” varied widely; nonetheless many governments and companies have meanwhile included “sustainability” in their policies and guidelines. Since 2003 a number of initiatives have been established to find consensus based approaches to considering sustainability in remediation in Europe, North America and Australia. Sustainable remediation is discussed in detail in Chapter 20 by Bardos et al., this book. This is to be expected today when issues such as world-wide urban pressures, water supplies, and climate change are key focuses across the political, social, economic and environmental sectors. These aim to define what “sustainable remediation” really means, in terms of what it describes and how the sustainable remediation projects will be implemented. The aim of NICOLE is to push this theory into practice. NICOLE members are professionals in the remediation sector and are subsequently very active in the on-going sustainable remediation debate. In 2008 NICOLE initiated a sustainable remediation working group (SRWG).

This group is defining what sustainable remediation principles actually mean, what tools are available, and importantly, which projects are really feasible. The SRWG is organised into several subgroups; case studies, economics, communication, Risk Assessment, sustainability and key performance indicators. These cover the critical aspects of sustainable remediation. The SRWG defines sustainable remediation (taking a view originally from SuRF UK[2]) as “a frame-work in order to embed balanced decision making in the selection of the strategy to address land and/or water contamination as an integral part of sustainable land use”.

The SRWG guidance will include a “roadmap” developed by the SRWG which sets out a stepwise approach to sustainability-based decision making for remediation, as shown in Fig. 24.3, and will be illustrated with case studies. The roadmap can be implemented into any remediation project, across projects and countries. The common principles in a simple road map are intended to improve the effectiveness of decision making. This NICOLE guidance will also include a check list of likely sustainable remediation indicators or metrics to help scope out the assessment, and a checklist of tools and techniques to help select assessment methods. The road map concept provides a scalable descriptive guidance about the key stages in achieving

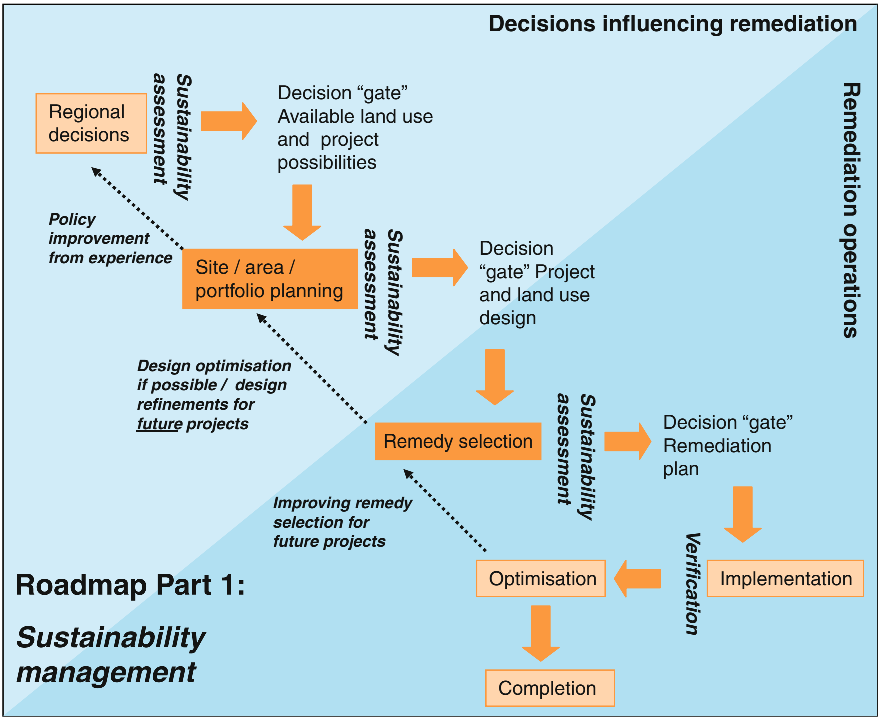


Fig. 24.3 Sustainable remediation management and assessment NICOLE road map – subject to revision (version January 2010)

---

[2] www.claire.co.uk/surfuk

sustainable remediation. It is not intended to be prescriptive as circumstances vary. The SRWG hope that this guidance will promote sustainable remediation widely through the European contaminated management sector. The guidance is available on www.nicole.org.

More details on sustainable remediation are discussed in Bardos et al. (Chapter 20 of this book).

24.10.2 Applied Sustainability – A Case Study

The former *Philips Site of Strijp S* in the city centre of Eindhoven is currently being redeveloped and transformed into a unique multi-use development for living, leisure and work, which will involve a combination of new buildings and landscaping and restoration of historic buildings around the site. During the development of the plans very ambitious criteria were formulated for sustainability. As a result, for the first time in Dutch history it was suggested that groundwater remediation should be combined with groundwater energy supply. However, the combination of sustainable energy and remediation and their objectives results in certain contradictions. Sustainable energy can be obtained from groundwater by pumping large flows and extracting heat or cold with a heat pump. The aim of a groundwater energy system is to maximise the energy capacity, which demands large groundwater flows. In the case of groundwater remediation, on the other hand, flows are usually kept to a minimum to meet the remediation approach (containment or load reduction in an agreed timeframe) and reduce costs..

A second paradox lies in the containment of groundwater. In a traditional approach with Heat-Cold-Storage, groundwater is pumped from a cold zone to a warm zone: using such an approach in a contaminated aquifer would potentially result in an increased movement and spread of the contamination. A remediation approach is primarily designed to contain and reduce the extent of contaminants, therefore the Heat-Cold-Storage approach apparently contradicts the goal of remediation. The solution that has enabled this approach to be successful required a change of the basic concept for the groundwater system. Instead of using cold and warm zones in the subsurface, it was decided to use a recirculation system. This system of extraction and infiltration wells uses a constant flow direction and extracts heat or cold from groundwater with a constant temperature (at the Eindhoven site 12–13 C), see Fig. 24.4. This system allows for energy extraction, while containing the flow of groundwater within a confined area, reducing the opportunity for the migration of contaminants. Moreover, this procedure stimulates natural degradation. The recirculation system leads to an increased mixing of the ingredients for natural degradation (contaminants, microbes and nutrients).

This synergy between the needs for Heat-Cold-Storage and groundwater remediation leads to significant reductions of CO_2 emissions, approximately 3,000 tonnes (50%) per year. The use of natural gas will decrease for this 27 ha site from 2.8 million cubic metres to less than 0.6 million cubic metres. At the same time, the

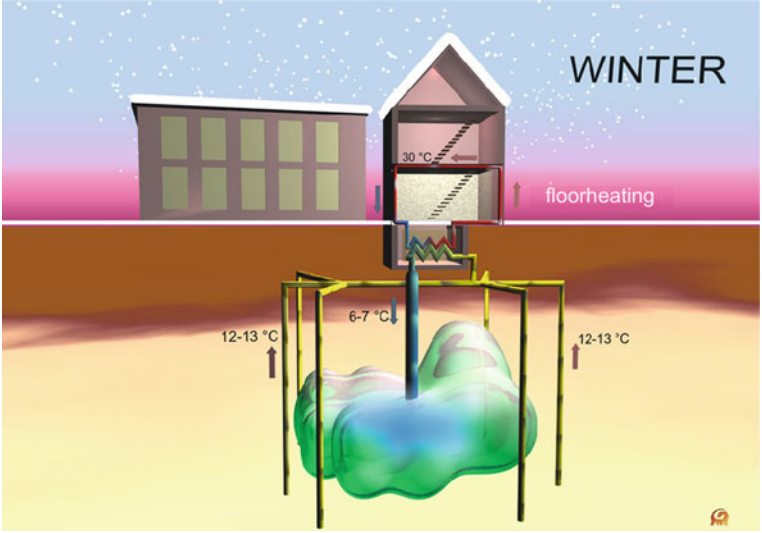


Fig. 24.4 Integrating Hot-cold storage with groundwater remediation at Eindhoven

use of electricity will increase from 2.4 to 4.7 million kWh as a result of the use of heat pumps. It can be concluded that this balancing will lead to a cost reduction of between 30 and 40%.

In the Netherlands there are hundreds of areas similar to *Strijp S*, which as a result of their specific character, dynamics and possibilities, challenge developers, owners and buyers to develop innovative, sustainable solutions. The Dutch government wants to transform the Netherlands into one of the cleanest and most energy efficient countries in Europe. Groundwater energy is an essential element in this strategy. The redevelopment of Strijp S provides sustainable energy in combination with an improvement of the soil quality.

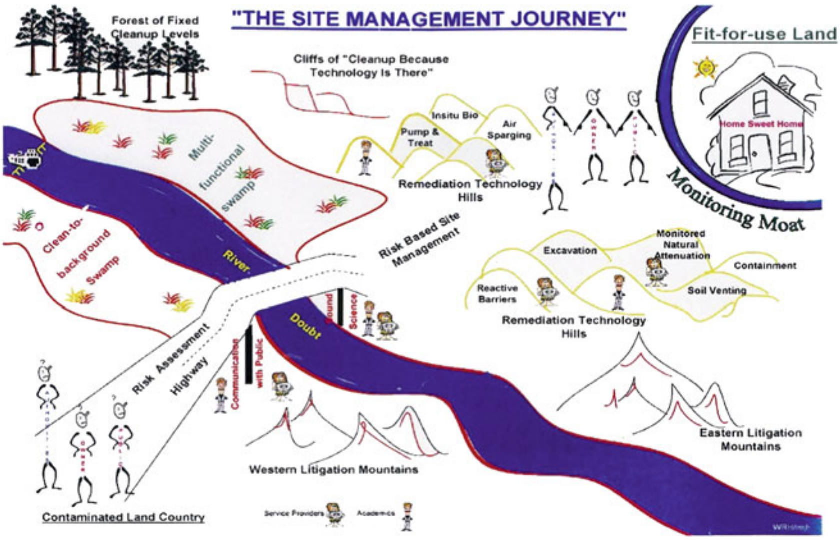
24.11 Conclusions

Since its formation in 1996, NICOLE has been at the forefront of developments in the contaminated sites remediation sector in Europe. Its members have developed a broad range of knowledge and experience in dealing with Brownfield sites, which have steadily increased in number in Europe, as many industrial sites have reached the end of their useful life. As stakeholders and problem holders in Brownfield remediation, NICOLE's industrial members have developed and put into practice a range of techniques and methods to improve, alleviate and prevent contamination of soil, surface water and groundwater. A greater knowledge and understanding of the forms and nature of contamination has allowed the refinement and improvement of remediation approaches and has developed the remediation sector to such an extent that the roadmap of contaminated sites remediation developed in 1998, designed as an illustration of the process, has been completely redeveloped within ten years. NICOLE and its members advocate the view that sustainable remediation, processes that mimic natural activity and a proportionate, risk-based approach are key elements of modern site remediation. It is not always necessary or practical to remediate a site so that it is completely free of contamination; the goal should be to remediate a site such that it becomes suitable for an onward new purpose.

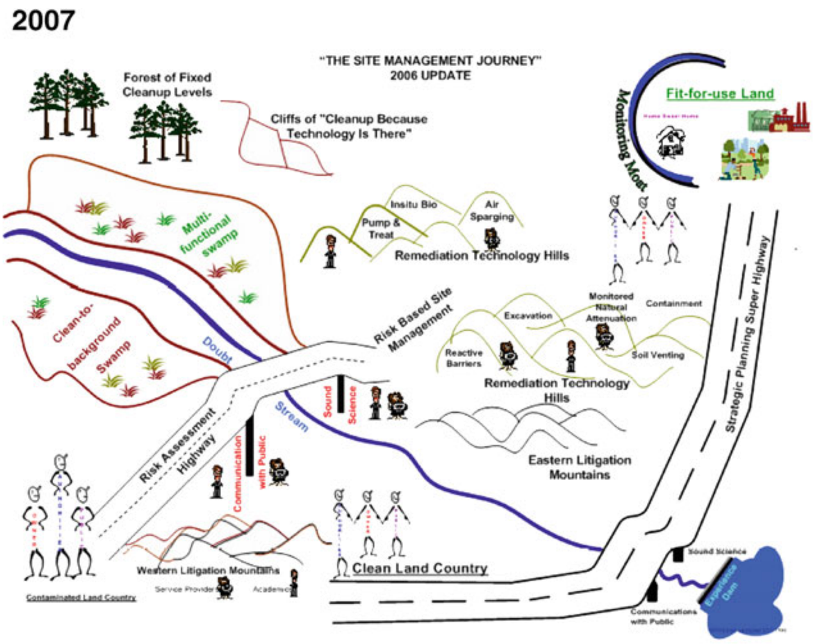
This chapter has illustrated the means by which this has been achieved using a range of examples and case studies provided by NICOLE industry sub-group (ISG) members. This success has been achieved against a background of developing environmental legislation that has increased in scope and complexity in the last 12 years. In many cases, NICOLE and its members have played an active role in the development of this legislation. NICOLE is now an established force in contaminated sites sector in Europe and it will continue to provide a forum for discussion, a source of advice and a political voice for industry.

Appendix: 10 Years of Progress: Two Road Maps to Contaminated Land Management

1998



Network for Industrially Contaminated Land in Europe. This pictogram was inspired by the CLARINET Steering Group discussion of the need to be able to map the decision making route that people face when confronting contaminated land issues. Taking into account the ideas discussed within CLARINET and NICOLE, this simple map was developed. It is not intended to show a single route from one point to another. Rather it is intended to show the various routes, challenges and obstacles that authorities, owners and the public jointly face when confronting a potentially contaminated piece of land, and how they might move together to having the sense of reaching a secure fit-for-use conclusion for its remediation and its present or future use. CLARINET is continuing to progress the goal of articulating the "decision making route" via development of a Strategy Paper on "Sustainable Contaminated Land Management: A Risk Based Land Management Approach".



What is Behind Redrawing the Site Management Map?

While at first glance the "Site Management Journey" 2007 Update may look very similar to the original Journey Map of 1998, there are actually quite a few changes that have appeared on the contaminated land landscape. It is hoped that the changing landscape continues until things like the swamps are no longer present, and the "Litigation Mountains" are almost eroded away. However, the changes between 1998 and 2007 are very positive:

The players (the little "stick-men" in the pictures	There is no change to the players. We had the right parties involved and need to keep them involved (land owners, authorities, public, service providers, and academics/researchers).
The "River Doubt"	Now dammed upstream by "Experience", and only a much smaller "Doubt Stream" continues to flow as a much less difficult to cross obstacle.
"Contaminated Land Country"	Becoming a smaller less populated place as sites are brought back into productive use.
"Clean Land Country"	A place where more and more of our activities begin by taking a more strategic approach to not creating new contaminated land.
The ferry route	The ferry that crossed the "River Doubt", between the shrinking "Clean-to-Background" and "Multifunctional" swamps, has gone out of business. Few people are any longer are mislead into trying to take that difficult and painful route through the swamps, and then facing the "Forest of Fixed Cleanup Levels" and the "Cleanup Because Technology is There Cliffs", to reach "Fit-For- Use Land".

The "Risk Assessment Highway"	The amount of traffic on the "Risk Assessment Highway" is going down as a result of the construction of the new modern "Strategic Planning Super Highway" from "Clean Land Country" directly to "Fit-For-Use" Land.
The new "Strategic Planning Super Highway"	Built directly from "Clean Land Country" to "Fit-For-Use Land". It entirely avoids the old "Risk Assessment Highway", and the off road "Risk Based Site Management" expedition through the "Remediation Technology Hills". The pillars holding up the new highway bridge are the same "Sound Science" and "Communications with Public" supports that successfully support the existing "Risk Assessment Highway."
"Fit-For-Use Land"	It is recognized that "Fit-For-Use Land" supports other valuable activities other than just "Home Sweet Home", such as industrial sites, parks, etc.
The "Litigation Mountains"	Have eroded somewhat, and are more hills now that pose less of an obstacle to successfully reaching "Fit-For Use Land".
The "Monitoring Moat"	The "Monitoring Moat" around "Fit-For-Use Land" is smaller since fewer sites require after care as they are remediated over time. The "Strategic Planning Super Highway" from "Clean Land Country" does not require a protective moat.

References

Carlon C, Swartjes F (2007) Analysis of variability and reasons of differences. In: Carlon (ed) Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation opportunities, JRC PUBSY 7123, HERACLES. European Commission Joint Research Centre, Ispra

NICOLE (2003) Report of the NICOLE workshop: management of contaminated land towards a sustainable future: opportunities, challenges and barriers for the sustainable management of contaminated land in Europe. 12–14 March 2003, Barcelona, Spain. Available on www.nicole.org. Land Contamin Reclam 11(4):449–472

NICOLE (2007) Report of the NICOLE workshop: redevelopment of sites – the industrial perspective . 14–15 June 2007, Akersloot, the Netherlands. http://www.nicole.org/publications/library.asp?listing=1. Accessed 25 August 2010

NICOLE (2008) Report of the NICOLE/SAGTA workshop: sustainable remediation 3rd March 2008, London, UK. NICOLE Secretariat, Deltares/TNO, Appeldoorn, the Netherlands. http://www.nicole.org/publications/library.asp?listing=1. Land Contamin Reclam 16(4): 381–403

NICOLE (2004) Report of the NICOLE workshop: sharing experiences in the management of megasites: towards a sustainable approach in land management of industrially contaminated areas. 29–31 October 2003, Lille, France http://www.nicole.org/publications/library.asp?listing=1. Land Contamin Reclam 12(2):127–158

Chapter 25
Sustainable Brownfield Regeneration

C. Paul Nathanail

Abstract Sustainable Brownfield regeneration involves making abandoned, underused, derelict and, only occasionally contaminated, land fit for a new long-term use in order to bring long-lasting life back to the land and the community it lies within. Brownfields sites have been affected by former uses of the site or surrounding land; are derelict or underused; are mainly in fully or partly developed urban areas; may have real or perceived contamination problems; and require intervention to bring them back to beneficial use. While Brownfields do not have to be contaminated, contaminated sites are the focus of this book so it is important to point out that risk based contaminated land management is an essential prerequisite to ensuring efficient deployment of resources to deliver land that is fit for use. Vision and strong leadership are needed to build up and maintain momentum during the long time for remediation, reclamation and redevelopment and before regeneration can begin. Brownfields occur throughout the world and, while local definitions of Brownfield may vary, there is growing consensus on the opportunity they offer and great benefit on sharing experiences of and good practice in their regeneration. Specialist Brownfield regeneration process managers are needed to help deliver more successful projects. Suitable enabling policy and facilitating public sector finance usually lag behind the structural change that causes Brownfields yet must respond quickly if regions are to survive and deliver the stability and opportunity their citizens have come to expect.

Contents

C.P. Nathanail (✉)
University of Nottingham and Land Quality Management Ltd, Nottingham, UK
e-mail: paul@lqm.co.uk

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6_25,

25.1 Doing the Right Thing – Right

Peter Druker defined management as doing things right and leadership as doing the right thing. This book is about the right way to manage contaminated sites. This chapter focuses on the intended reuse of “unemployed” land that may or may not have been affected by contamination. Sustainable Brownfield regeneration is the right thing to do. So anyone who is reading this chapter deserves either congratulations for reaching this far in a book on doing things right or an invitation to consider this chapter’s topic as the end point of much, but not all, contaminated site management and therefore to return to earlier chapters to appreciate how such management should be done.

Growing realisation of the need to conserve and target resource consumption has resulted in widespread agreement that land should not be consumed or sterilised by one land use thereby preventing subsequent land uses. Past industrial and waste management practices have however created conditions which at least hinder if not prevent future land uses. The earlier chapters in this book deal with assessing and mitigating the risks associated with chemical contamination. Risk Assessment must precede remediation if remediation is to deliver tangible benefits to offset its cost and ecological footprint.

25.2 What are Brownfields?

Europe’s sustainable *Brownfield regeneration network*, CABERNET (www.cabernet.org.uk), defined Brownfields as sites which:

- have been affected by former uses of the site or surrounding land;
- are derelict or underused;
- are mainly in fully or partly developed urban areas;
- may have real or perceived contamination problems; and
- require intervention to bring them back to beneficial use (CABERNET 2006).

For the purposes of this book, it is important to realise that Brownfields are not necessarily contaminated, and may never even have been subject to a potentially contaminative land use.

The term Brownfield has developed various legal, formal and colloquial meanings in different, and occasionally within individual, countries. It is best understood in terms of a candidate redevelopment site as the opposite of a Greenfield site. Very densely populated countries such as the UK and the Netherlands do not have the luxury of discarding land and relying on urban sprawl to provide new land for development. Instead such countries have had to develop robust policies and technologies to ensure they reuse land once a given land use has come to an end and thereby prevent urban sprawl and protect as yet undeveloped "Greenfield" land. Brownfield redevelopment may have to be, but is not necessarily, preceded by remediation to reduce contamination related risks, the subject of this book, and by reclamation from unsuitable conditions such as inundation. However, the objective of the construction of new buildings and infrastructure on a site is usually to achieve the return of some form of human presence and activity, be it for living, working, retail, manufacture or leisure.

Brownfield sites include:

- closed petrol stations;
- former military bases;
- disused warehouses;
- derelict office blocks;
- dilapidated housing;
- abandoned offices;
- disregarded monuments;
- discarded railway lands;
- infilled landfills.

And some examples of what are NOT Brownfields include:

- operating factories;
- occupied houses;
- construction sites;
- in use military ranges;
- farmland, whether or not affected by pesticides and fertilizers;
- overgrown allotments;
- burnt out forest areas;
- recolonised colliery spoil heaps.

25.3 What is Regeneration?

Regeneration is "the process of turning round deprived communities in decayed neighbourhoods". Regeneration is the motivation of most redevelopment. Ireland's new O2 music venue (Fig. 25.1) is billing itself on hoardings in Dublin Airport: "It's built – they're coming." "Build and they will come" was the mantra of the Kevin Costner film "Field of Dreams", but it was the US EPA's Carl Alvarez who introduced the phrase to a European Brownfield regeneration audience at the first CABERNET plenary meeting in Athens in January 2003. He pointed out that for every dollar of public sector investment, four to five private dollars followed. This ratio has also been observed in the UK and other parts of Europe.

Fig. 25.1 Promoting the new Dublin O2 (source: C P Nathanail, reproduced with permission)

25.4 What is Sustainable Regeneration?

> *a time to be born and a time to die,*
> *a time to plant and a time to uproot,*
> *a time to kill and a time to heal,*
> *a time to tear down and a time to build,*
>
> (Book of Ecclesiastes Chapter 3 verses 2 and 3)

Nothing lasts forever. Regeneration involves reinvention and long-term creation, or at least maintenance, of capital: economic, intellectual, social and environmental capital. Many conflate and confuse sustainable development and sustainability. The Brundtland Commission definition of sustainable development has taken on the mantle of a mantra: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs." Not only does the debate confuse "wants" with "needs", but also "growth" with "development". The laws of compound interest seem to have been forgotten or cast aside to enable economic growth to be a necessary precondition for what is termed sustainable development in much of the developed world.

In many senses sustainability is a journey rather than a destination, a process rather than a product. Some cities have achieved millennia of longevity, Athens, Alexandria, Beijing, Jerusalem, Rome; some mere centuries: Constantinople, New York, London, Boston, Melbourne; still others are new born: Brasilia, Canberra. Some have come and gone: Thebes, where the Sphinx's riddle led to many a painful death, Nineveh, Jonah's destination after his Piscean sojourn, Amarna, monotheist Pharaoh Akhenaton's capital, the legendary Atlantis. At some point, demolition debris acquires the status of archaeological find: rubbish becomes artefact. Rarely cities are built from scratch, and even more rarely they are built to last: Masdar in Abu Dhabia is to be designed by Mott MacDonald as the world's first zero waste, carbon neutral city; meanwhile on China's third largest island, Arup are designing the Dongtan Eco-city.

London was born due its location: it is one of the first places the Thames could be crossed. It grew to become the capital city of England and the United Kingdom and one of the few truly world cities. From its Roman foundations London has been reclaimed and redeveloped many times as its communities regenerate and flourish (Fig. 25.2).

Tower 42, London's first sky scraper, was built for the NatWest Bank but had to be *re*clad and *re*furbished after it was damaged in 1993. Names clearly matter: 30 St Mary Axe also goes by the moniker Swiss Re Building, but is best known by its metaphorical nickname, the London Gherkin. Its height is deceptive: at 180 m it is three times as high as Niagara Falls (http://www.30stmaryaxe.com/). "St Paul's is a lasting monument to the glory of God and a symbol of the hope, resilience and

Fig. 25.2 Reflections on London old and new: Tower 42 (the former Natwest Tower), Swiss re building (aka the London Gherkin) and St Paul's Cathedral mirrored in office window (source: C P Nathanail, reproduced with permission)

strength of the city of London and the United Kingdom" (www.stpauls.co.uk). A cathedral has stood in east London 350 m north of the river Thames since 604 AD. The predecessor of today's building was destroyed in the great fire of London in 1666. Sir Christopher Wren used the space created by the devastation to *re*define the built environment for centuries. Its environs have not benefitted from such longevity. Paternoster Square emerged from the bombed ruins of World War II: "[By 1967], the area around St Paul's had been rebuilt. This was not exactly the broad, sunlit uplands post-war urban planning might have been; more the grim, windswept plazas of contemporary Fleet Street cliche. Yet at least the glum offices offered no threat to Wren's monument" (Glancey 2003). In 2003 Glancey also reported "everything has changed. After a protracted struggle, Paternoster Square Mk2 is finally complete – and St Paul's is now flanked to its north by a gathering of burly office blocks clad in the architectural equivalent of tweed coats. The odd bit of classical paste and some bizarre 1930s Italian fascist-style posturing help complete the look of this architectural fancy dress party".

When the Persian fleet threatened Athens, the oracle at Delphi claimed the Athenians defence lay in a "wall of wood": the navy. And so it proved: small, highly manoeuvrable Greek triremes destroyed the numerically superior, larger more cumbersome Persian warships in the confined waters off Salamis. What began on land with the 300 at Thermopylae and was to finish with the good news from Marathon received essential momentum by Themistocles' use of his wall of wood. Silver paid for the ships; silver won from the mines of Lavrion (Fig. 25.3) at a price that its community would continue to pay for the next two and a half millennia as lead in the ore damaged children's healthy development. Yet, Lavrion survived and so did the Athenian seeds of western civilisation.

The work of the CABERNET network highlighted some high level principles that seem to characterise regeneration that is generally thought to have been effective, lasting and who knows, in the long run even sustainable (Table 25.1). While these

Fig. 25.3 The ruins of the Lavrion silver mine in Attica, Greece (source: C P Nathanail, reproduced with permission)

Table 25.1 Seven principles of effective regeneration

Principle	Comment
People matter – absolutely	The link between environment and human health in its broadest sense of well being requires local concerns to be taken into account during the, potentially long term, redevelopment phase
"Places for people"	Design for people rather than minimum cost or maximum return on investment
Having a shared Vision is vital	Regeneration involves long time frames and often the price & inconvenience is borne by one group or generation but the benefits reaped by another
There is no *I* in team	Regeneration is a complex sequence of processes and involves many people who must work together towards a common end
Build and they will come	Doing nothing is not an option – merely a sure fire way of things getting worse
Waste is a resource in the wrong place	Creative thinking can reduce waste generation and turn redundant materials and buildings into useful resources
Leaders serve others now and in the future	The challenge to leaders is to ensure a servant attitude towards those who will benefit from their efforts – in both the medium and long term

are not necessarily relevant on every project or may have different relative importance in different projects they do form a useful core on to which a project specific vision can be crystallised.

25.5 Re Concepts in Regeneration

Each English region has its own Regional Development Agency (RDA). Advantage West Midlands (AWM) is the RDA for the area that includes England's second, and now largely post industrial, city: Birmingham and its Black Country hinterland. RegenWM is funded by AWM to promote and develop regeneration excellence in the West Midlands (Regen 2009).

At a series of workshops on sustainable regeneration hosted by RegenWM – delegates were invited to list as many words beginning with the prefix "re" that had some *re*levance to Brownfield *re*mediation, *re*clamation, *re*development and *re*generation. Does the difference between these terms matter – not *re*ally, does it? Well, as the old adage goes, if you don't know where you are going, how will you know when you have got there? Or even know that you are on the right path?

Remediation, as parts of this book ably explain, involves demonstrably breaking the source-pathway-receptor contaminant linkage,[1] thereby reducing risks to below a predetermined level (Nathanail and Bardos 2004).

---

[1] This book uses the term "contaminant linkage" to refer to the combination of source-pathway-receptor coned by the UK and termed a pollutant linkage. For ease of cross reference and consistency this chapter has adopted the term 'contaminant linkage' but it is not a term widely found in the literature.

Reclamation involves reclaiming land, bringing it back from a physical condition in which it could not be used into a condition where it can be built on or otherwise be put to beneficial use.

Redevelopment takes remediated and/or reclaimed land and creates new landscape, buildings and infrastructure.

Finally, regeneration happens when people come to the new development, bringing with them life and activity. One activity may fail and another take its place with no need for redevelopment, though perhaps some *re*furbishment or *re*fitting may be necessary. A classic example of such a change in function is the Millennium Dome, London's second most popular paying visitor attraction during 2000 and now, as the *re*branded O2, the world's most popular music venue.

The workshop delegates identified well over 60 *re* terms. While many are used interchangeably, this should be avoided as it confuses and diminishes the rich variety of activities and processes that Brownfield regeneration entails.

25.6 Brownfield Regeneration: A Multi Stakeholder Challenge

Taking a former gasworks, derelict apartments or out of date retail parks and creating a vibrant place, requires a combination of creative, scientific, engineering and other skills coordinated by a Brownfield regeneration process manager. CABERNET recognised this multi-stakeholder aspect of Brownfield regeneration and assembled its members into eight groups.

Such a multi disciplinary approach to urban land management requires personnel with special skills and knowledge base. While specialism is important, the willingness and ability to communicate with other professionals to achieve an optimal solution to the overall problem must be based on mutual respect and understanding of different perspectives. A partial or blinkered view may well compromise overall effectiveness for the sake of maximising one specific aspect. For example focusing on contamination risk mitigation may result damage to ecosystems, prolong a project or increase the budget to the point of non viability.

25.7 The CABERNET Brownfield Process Manager

Project management is an established discipline. However, its application to Brownfield regeneration requires additional elements of visioneering, a holistic appreciation of different, often conflicting, stakeholder perspectives and a commitment to "making it happen" (CABERNET 2006; RESCUE 2005).

Sustainable Brownfield regeneration requires individuals who understand at a strategic level the part each professional can play in the process and most importantly who have the ability to identify the opportunities Brownfield sites can represent. Brownfield process managers need to look top-down at a Brownfield site and its community, in order to place the site in the context of the community through development of what CABERNET termed an "opportunity plan" which focuses on project specific opportunities. However, they also need to follow a bottom-up approach that builds on the site and community's strengths and minimises or mitigates any weaknesses. As CABERNET's professional skills working group, led by Francesca Neonato (PN Studio, Italy) and Euan Hall (Land Restoration Trust, UK) pointed out: such individuals exist today by dint of hard won experience but:

> Whilst many environmental/land based professions encourage continuing professional development (CPD) in technical areas, few concentrate or give much regard to managerial leadership and process-oriented skills. There is a need for a European wide recognised post-graduate diploma or certificate which builds on technical knowledge/skills through enhancing leadership skills, consensus building for the range of skills potentially required by a Brownfield Process Manager (CABERNET 2005).

In essence, Brownfield process management involves applying the seven principles listed in Table 25.1 in ways that are appropriate to the sites in question. While technical competence is needed in project delivery, too strong an emphasis on technical issues can result in "the wrong thing being done right". Only by looking at the bigger picture can the "right thing" be identified and only then should the technical skills be deployed to "do the right thing, right". Brownfield process managers must both implement and influence policy and must both muster and master a wide spectrum of technical specialisms. They should also provide long term continuity to ensure project creep does not result in delivering something unexpectedly different from that envisioned and agreed. CABERNET (2005) illustrated the role of the opportunity plan (see Box 25.1) with examples from the work of CABERNET members across Europe and described the personal skills and attributes of a Brownfield process manager.

Box 25.1 The CABERNET Opportunity Plan (CABERNET 2005)

Jumping straight to a Master Plan (development framework) can result in unimaginative and inappropriate regeneration. A preliminary options appraisal stage is suggested to define the range of regeneration opportunities for the site. However, these opportunities should be anchored in and reflect the site environmental, economic and social setting/ context – an opportunity plan. The spatial extent to be considered will be that impacted on and impacted by the regenerated site. This "zone of influence" varies from site to site.

The opportunity plan should not be an immensely detailed piece of work replacing a masterplan, but an informed review of where Brownfield land sits in relation to existing communities – its spatial context. From this, an appropriately skilled professional can then determine where development should best take place if there are a variety of options…

The Opportunity Plan will allow the BPM to identify a series of potential regeneration options (called opportunities) from this understanding of the site and its context. The opportunity plan allows the Brownfield process manager (BPM) to put that site in the context of the settlement in which it is located and explore the various needs for redevelopment. It comprises a spatial strengths-weaknesses-opportunities-threats SWOT analysis of the site in its current and likely future context which spawns a series of feasible regeneration options. . .

So what are the different facets of the assessment – the SWOT analysis.

- Strengths – may include skilled workforce, diverse economic base, location, age of workforce, cultural heritage etc.
- Weaknesses – unskilled workforce, single industry, poor locations, long term health problems etc.
- Opportunities – new business, community spirit, skilled workforce, location etc.
- Threats – dereliction, crime, health, unemployment, displacement of services, jobs etc. to other settlements.

As can be seen the Opportunity Plan requires a Brownfield Manager to undertake a broad analysis, not just a technical site assessment, if benefits are to be maximised. The technical assessment and detailed studies will follow on.

25.8 International Brownfield Definitions

There is no internationally agreed definition of the term "Brownfield". Three definitions from the EU, UK and USA are discussed below. The consequence of these differing definitions is that specific sites considered Brownfields under one definition would not be under another (Table 25.2). The consensus view is that Brownfield sites pose obstacles to their redevelopment in addition to those posed by previously undeveloped sites. What is less widely appreciated is that Brownfields can come with a *de* facto dowry of infra-structure or good access or location or previous workforce that would be difficult to create from scratch on a "Greenfield" site.

25.8.1 Europe Union

In the mid 1990s, CARACAS (Concerted Action on Risk Assessment for Contaminated Sites in the European Union; Umweltbundesamt 2009) brought together scientists, regulators and policy makers to define the state of the art of contaminated land Risk Assessment across Europe (Ferguson and Kasamas 1999) and to develop a consensus on what such Risk Assessment involved (Ferguson et al. 1998). CLARINET (Contaminated Land Rehabilitation Network for Environmental Technologies in Europe) extended the work of CARACAS to consider how to rehabilitate contaminated sites. This included considering Brownfields and led to the formation of the CABERNET network and the RESCUE project.

Table 25.2 Comparison of Brownfield definitions

	Land use			Brownfield status		
Site	Intended	Current	Photograph	Europe	UK	USA
Former furniture store, Nottingham UK	Derelict	Place of worship		No	Yes	No
Former landfill, Los Angeles USA	Derelict, contaminated, geotechnically unstable	Mixed commercial and retail		Yes	Yes	Yes

Table 25.2 (continued)

Site	Land use		Photograph	Brownfield status		
	Intended	Current		Europe	UK	USA
Former metaliferous mine, Avoca EIRE	Derelict and locally contaminated			Yes	No	Yes
Spinalonga Fortress, off Crete GREECE	Monument	Monument		No	No	No

Source of all photographs: C P Nathanail, reproduced with permission

There is no common definition of Brownfield across the 27 member states of the European Union. However two have gained widespread acceptance over the last decade. The EU funded CLARINET network produced a definition which was subtly but significantly modified by CABERNET (2006) to include sites that:

- have been affected by former uses of the site or surrounding land;
- are derelict or underused;
- are mainly in fully or partly developed urban areas;
- *may* have real or perceived contamination problems; and
- require intervention to bring them back to beneficial use.

The effect of CABERNET's three letter modifier "may", in italics above, is to remove the *necessity* of contamination related issues for a site to be considered a Brownfield. For example whole streets of abandoned miners' cottages in villages in the north of Nottinghamshire or of ornately decorated town houses in Riga would usually be considered as Brownfields, but are unlikely to be affected by contamination issues.

As early as 2004, the failure of the EU to achieve the Lisbon targets of 2000 was reported by the BBC as a medium to long term threat to the sustainability of the society Europe has built (http://news.bbc.co.uk/1/hi/business/3979619.stm). There was high level recognition that "Each element of the Lisbon strategy is still needed for the success of the whole. Improved economic growth and increased employment provide the means to sustain social cohesion and environmental sustainability. In their turn, social cohesion and environmental sustainability can contribute to a higher growth and employment" (Kok 2004). However, the Kok report warning that "Europe, in short, must focus on (economic) growth and employment in order to achieve the Lisbon ambitions" is flawed, as any student of compound interest can work out, unless such growth can be decoupled from environmental impact and from contributing to social inequality in other parts of the world. Kok's report makes no mention of Brownfields or even of land. Such a spatially-blind analysis reflects the weakness or inability of the EU to recognise the territorial element in its socio-economic venture. A more recent, but still pre-credit crunch, analysis was more optimistic and highlighted the potential role of eco-efficiency in driving the European economic engine.

Whatever the future holds for Europe, the way it manages its land will be central to the long term maintenance of peace and stability on a continent that has only recently begun to put behind it a millennium of land-grab related conflict. Smart reuse of the sites of former factories, mines and electronic assembly plants is an essential pre requisite of social equity and wise environmental stewardship.

25.8.2 UK

The UK has had a long history of redeveloping former industrial sites. It now enjoys not one but two euphemistic meanings of the "Brownfield" term, including one policy based definition. England has a policy that 60% of new housing will be built on “previously developed land” (PDL), which is commonly referred to as “Brownfield”. For example, English Partnerships (now part of the Housing and Communities Agency) is custodian of the National Brownfield Policy that focuses on returning PDL to beneficial use. Much post industrial land affected by contamination has been redeveloped for housing. Such “chemically challenged” land is also often referred to as Brownfield. The 60% target has been met, indeed exceeded, only if the broad definition of Brownfield (aka PDL) is used. However an entire section of the consultancy sector prefers to refer to land affected by contamination as “Brownfield”: c.f. the nonsensical terms Brownfield remediation. This dichotomy was recognised and became the subject of several debates in the House of Commons (Box 25.2). Such debates demonstrate the political and economic sensitivity around Brownfield definitions which are used to target financial or other interventions.

Box 25.2 Extract from Parliamentary Debate on the Definition of Previously Developed Land

From Hansard (2008)
Gardens
Mr. Burstow: To ask the Secretary of State for Communities and Local Government when she expects
2 Apr 2009: Column 1504 W
the review of the Brownfield designation of garden land to be completed; and if she will make a statement. [267227]

Margaret Beckett: During the progress of the Planning Act 2008, the Government committed to review the evidence on the extent and impact of housing development on garden land in order to establish whether there was a genuine problem. This commitment was undertaken in the absence of firm evidence that a problem existed.

We propose to establish the amount of housing development on garden land, which at present cannot be distinguished from other land classified as “previously-residential” in the Land Use Change Statistics, such as estate regeneration or conversions.

The review will take place in two stages. The first stage will involve obtaining data directly from local planning authorities and the Planning Inspectorate. My Department will write to local planning authorities in England requesting data on the number of planning permissions granted or refused (including the outcomes of any associated appeals), from 1 April 2003 to 31 March 2008, for housing development within the curtilage of an existing dwelling house – but only where these permissions have, or will, result in a net increase in dwellings within the existing curtilage.

We have worked with the Planning Officers’ Society and the Central Local Information Partnership in designing the process of data collection for the review to ensure it as straightforward for local authorities as possible.

We will exclude from the review details of regeneration or renewal schemes, as these could seriously distort the data. We will also exclude extensions to properties, as these do not lead to an increase in dwellings, and conversions, as these do not generally lead to a loss of land.

Authorities will also be asked to provide a commentary on the development plan policies they rely on to determine these type of applications, and information on whether they have identified gardens (either individually or through the identification of a broad location), through the plan led system, for future development.

Local authorities will also be given the opportunity to say whether garden development is or is not an issue for them in their area and why.

As part of the first stage we will also ask the Planning Inspectorate to provide information on the appeal decisions highlighted by planning authorities, to help establish the principal reasons why the decisions were granted.

For the second stage of the review, we will seek external analysis of the data collected from the first stage to determine the impact of the development. This will probably focus on a subset of authorities which will need to include authorities from each region, and an even mix of large and small authorities in both urban and rural areas. We will publish further details of the second stage of the review at the time. Subject to the response rate we receive to stage one of the review, our aim is to conclude the second stage by the summer.

The purpose of the review is to ascertain if there is clear and genuine problem with the extent of development on gardens. We are committed to considering action if the evidence discloses a problem, provided that any changes should not have the effect of undermining our objectives on housing.

2 Apr 2009 : Column 1505 W

We shall also be writing separately to Opposition spokespersons and hon. Members in England to bring the details of the review to their attention.

25.8.3 USA

The US "Small Business Liability Relief and Brownfields Revitalization Act" uses the term Brownfield to mean "real (estate) property, the expansion, redevelopment, or reuse of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant". Each State determines whether or not individual sites meet this definition and therefore qualify for financial support for their remediation or not.

President Obama's economic stimulus package includes large sums for Brownfield sites. The American Recovery and Reinvestment Act of 2009 (ARRA) provided some $787 billion of which $100 million was allocated to the US EPA Brownfields Program for clean up, revitalization, and sustainable reuse of contaminated properties. Hot debates revolve around the classification of sites as Brownfield, so that they can qualify for support under this stimulus cash.

25.8.4 Comparison of Brownfield Definitions

The previous section illustrated some of the differences between Brownfield definitions. In many countries funding is allocated to sites that meet the relevant definition and therefore the implications of choosing alternative wording is important. Four sites are used to illustrate these differences (Table 25.2).

Former retail and warehouse premises near Nottingham Castle were abandoned when the furniture retail company that was leasing the buildings went into liquidation at the end of 2008. The site is currently derelict and the buildings unoccupied. A local church which had been hiring rooms in a secondary school has now bought the land with a view to either refurbishing the existing buildings or erecting a purpose-built place of worship.

A former landfill in Los Angeles is being redeveloped for mixed commercial and residential land use after remaining derelict for many years. The location of the site and its very good connection to the freeway network meant the costs of remediation and reclamation could be recovered in a short enough period of time to make the project viable.

The Avoca mine in the Republic of Ireland has been derelict since mining finished in the 1980s. It is now being remediated and restored to a mine heritage and recreational facility.

Spinalonga is a small fortress island off the coast of Crete in the Mediterranean. Its former uses include the Venetian fortifications and a leper colony. While unoccupied since the departure of its last resident, a Greek Orthodox priest, in 1962 the island is one of Crete's most popular tourist attractions.

The first three of these sites fall within one or more of the definitions of Brownfields described above. Spinalonga alone does not merely by dint of its continued "use" as an ancient monument.

What has become clear in the last decade is that few if any countries are immune from the need to consider their Brownfield legacy. Most of Europe's Brownfields have been created as a result of structural or geopolitical change. As large industries shut down or relocated, they leave behind a legacy of Brownfield sites. However even the booming economy of China has resulted in Brownfield sites and policies are beginning to emerge to return these to beneficial use. One advantage that a growing economy offers of course is economic capital to invest into Brownfield sites. Some Chinese Brownfields are early generation factories built in the 1950s and 1960s, around the centre of what were then relatively small settlements. As economies and populations grew, technological developments have rendered those factories redundant but surrounded by residential areas. Such Brownfields are now seen as highly desirable tracts of land onto which to develop residential, office and retail developments to service an increasingly affluent and urban populace.

Derelict, abandoned and underused land is a growing problem throughout Europe and beyond. Changes in industrial practice and in the geopolitical make-up of

Europe have resulted in large tracts of redundant land. The transition to an industrial economy took place at different times in different parts of Europe. In Britain, industrialization took place from the late 18th century to the onset of World War I. In Germany, the main period of industrialisation was from approximately 1870 until the onset of World War II. In Eastern Europe on the other hand, industrialisation took place from about 1900 until the early seventies. In the post industrial conditions that now prevail across most of Europe, many industrial sites have been abandoned as industries disappear (e.g., manufactured town gas) or move to countries with lower labour costs (e.g., coal mining). Uncontaminated or Greenfield lands (note that the two are not synonymous) have become a limited and on occasion scarce resource in the densely populated and highly industrialised parts of Europe. The reuse of the abandoned industrial sites is therefore essential. Indeed the effects of the "end of the noughties" credit crunch has been widespread closure of factories creating a fresh injection of Brownfield sites.

Alker et al. (2000) recognised the importance of a precise definition for such Brownfield sites: Policy was being written to deal with them and budgets allocated to facilitate their return to beneficial use. Whatever the definition, Alker et al. considered that Brownfields have a set of shared characteristics:

- breakdown of economics;
- problems in attracting new investors;
- high unemployment rate;
- adverse effects on urban life;
- decline of tax income for the communities;
- social conflicts;
- consumption of Greenfields.

Of course as we have seen in the discussion on Chinese Brownfields above, Brownfields occur in countries where the economy is booming. The city of Ningbo provides a good example where former factory sites now lie derelict and awaiting redevelopment in a city growing in size, with a successful economy and no significant unemployment problem.

Moreover, Brownfields are a consequence of a set of, perhaps unconscious, shared values that land is a resource to be consumed rather than managed. Indeed one global mobile phone company's attitude to real estate is to "pay as you go": that is to pay, at a reasonable premium, for assembly, office or retail space for as long as you want it and once you have achieved your local goals and objectives to be able to simply migrate from that location to the next one. This is reflected in land use planning policies and practice that focuses on the next land use rather than the entire life cycle of land use from construction to decommissioning. Such a consume and discard attitude may be understandable, but not condonable, when dealing with consumer products such as mobile telephones or clothes. However, even in low population density countries or regions such a consume and discard approach to land use has been shown to be flawed: the social consequences of the "left behind" population are externalised and often have to be picked up by the national tax payer or international aid agencies.

25.9 Typologies of Brownfield Sites

A typology is a way of describing a characteristic of a phenomenon: in this case Brownfield sites. The consequences of different definitions is illustrated in Table 25.2.

25.9.1 Economic

CLARINET (Ferber and Grimski 2002) proposed a three fold classification of Brownfield sites depending on the profitability of regeneration (Fig. 25.4). CABERNET (2006) developed the concept and analysed the role of private and public sector finance (Fig. 25.5). The ratio of reclamation costs and site value defines three types of sites. Where a profit is still to be made ("Type A sites") the private sector will undertake the reclamation process in a conventional investment project. In many situations, however, private developers and financiers are not able, or willing, to act on their own to ensure that the full economic potential of site reuse will be achieved ("Type B sites"). This means that Brownfield properties are often underused, because the perceived costs and risks of bringing them back into use exceeds the benefits to their owners. Thus some kind of public sector intervention is necessary to bring the land back into productive use (Alker et al. 2000). Where the reclamation costs are broadly equal to the site value, or uplift in site value, then public-private partnerships or the creation of infrastructure by the public sector can attract private investment to a site. Sites where reclamation costs are much higher than site value cannot reasonably be brought back in to economically beneficial use (Type C sites). However, such sites can be reclaimed for the public good.

> The need of quiet, the need of air, the need of exercise, and... the sight of sky and of things growing seem human needs, common to all men. Octavia Hill (co-founder National Trust).

The Land Restoration Trust (LRT) (2009) recognises that "many physical regeneration strategies prioritise regeneration for economic use. Whilst this is important it means that often land not suitable for commercial or industrial

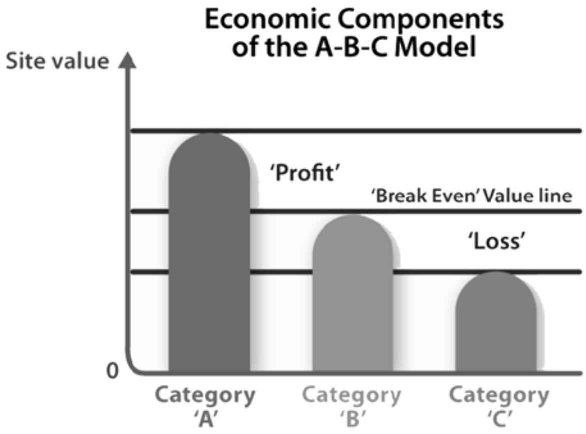


Fig. 25.4 CABERNET A-B-C model: The economic components of Brownfield sites (source: CABERNET 2006, reproduced with permission)

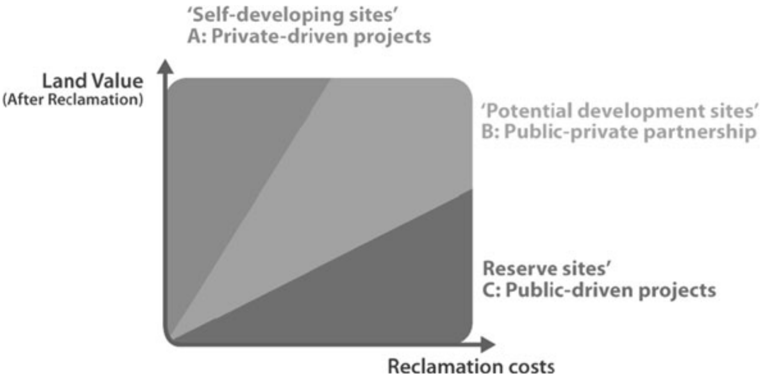


Fig. 25.5 CABERNET A-B-C model: The role of private and public sector finance (source: CABERNET 2006, reproduced with permission)

redevelopment ("hard end use") can be left derelict and under-used". The LRT is modelled on the National Trust which was founded in 1895 following concerns about the impact of uncontrolled development and industrialisation "to act as a guardian for the nation in the acquisition and protection of threatened coastline, countryside and buildings" (National Trust 2009). The LRT aims to acquire, reclaim and manage in perpetuity for the public good Type C sites.

25.9.2 *Temporal*

Sites may meet the CABERNET definition for short or longer periods of time. Their impact on their communities will grow the longer they remain in a Brownfield state. Policy makers have recognised this and begun to use the length of time a site has been Brownfield to prioritise intervention.

CABERNET (2006) likened the process of Brownfield creation and reuse to the filling and emptying of a bathtub (Fig. 25.6). Structural change creates Brownfields:

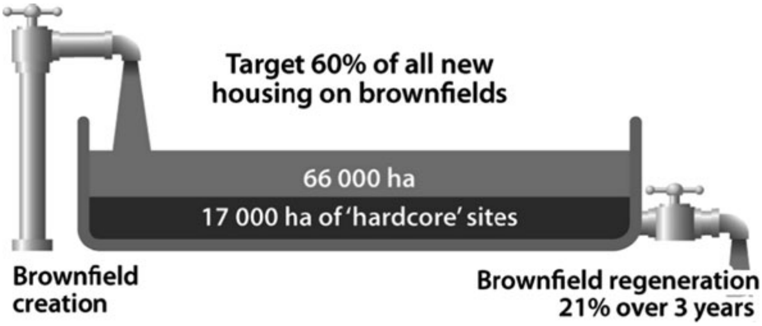


Fig. 25.6 CABERNET bathtub model: England's situation (source: CABERNET (2006), reproduced with permission) NB English policy relates to previously developed land which is commonly and perhaps euphemistically referred to as Brownfield land

the bathtub fills up. Private sector redevelopment and public sector intervention help return Brownfields to beneficial use: the bathtub is drained. However, some sites stay in the Brownfield state for long periods and become progressively harder to return to beneficial use: the sludge at the bottom of the bathtub. Such long term sites are the focus of national and regional programmes. If large enough, their impact on local communities can begin to affect successive generations and create chronic social problems that can become progressively harder to solve.

25.10 Sustainable Regeneration

Sustainable regeneration is a journey, not a destination. So sustainable regeneration must ensure that each cycle of land use does not end up in a *cul de sac*. In Bardos et al. (Chapter 20 of this book) the process of reducing unacceptable levels of soil contamination related risk to acceptable or tolerable levels is discussed. The processes of land reclamation, site redevelopment, building renovation, refurbishment or reuse are discussed by RESCUE (2005).

The journey of sustainable regeneration may involve several cycles of land reclamation, redevelopment or refurbishment. As cultures change, technologies emerge or disappear, fashions alter or indeed climates change, once necessary land uses become redundant and the land they occupied can be put to some new use. This land reuse is an essential component of successful urban land management and helps prevent both urban sprawl and derelict or underused urban cores.

The former car manufacturing capital of the world, Detroit, exemplifies the malaise of "unemployed" post industrial land. In May 2008, the US EPA (United

Fig. 25.7 Detroit suburbs seen from the air. Note the uniform housing developments creating monotone land use patterns that require car access to retail, education, work and leisure facilities (source: C P Nathanail, reproduced with permission)

States Environmental Protection Agency) held its biennial Brownfields conference in Detroit. At the time, Downtown Detroit had many empty downtown properties and traffic free streets. Yet the sprawling suburbs were characterised by low density housing on a rectilinear grid (Fig. 25.7).

Long term success in regeneration is possible. Cities such as Bilbao, Manchester, Glasgow and even war ravaged Beirut have shown it is possible to reinvent and reattract. Incremental progress must sometimes be accompanied by step-changes whether externally imposed (e.g., Manchester or Detroit) or internally engineered (e.g. Glasgow). However, what the regeneration process in many successful cities seem to have in common is vision.

25.11 The Need for Vision

Failing to plan is planning to fail

This chapter began with the recognition that right regeneration is regeneration that lasts. However urban land management is a slow process: the champions of the London 2012 Olympics have asked to be judged in 2050 when the fruits of their regeneration efforts in East London can be properly evaluated. Canary Wharf took two owners to bankruptcy before it became a lynchpin of the world's finance system. The O2 emerged phoenix-like from the ashes of the reputation of the millennium dome.

Foxell (2008, p. 235) reports that much of Richard Rogers 1986 "London as it could be" master plan for a Royal Academy exhibition "have become reality". Rogers (2007) "felt great opportunities to improve the capital were being ignored in favour of a piecemeal approach to planning, led by market forces rather than by any consideration of the wider public interest". Roger's vision revolved around the twin axes of the Thames Embankment from Westminster to Blackfriars and "the route across the Thames from Waterloo station (already projected as the terminus of the new Channel Tunnel rail link) to Trafalgar Square."[2]

CABERNET recognised that Brownfields provide essential room for manoeuvre in urban land management (Fig. 25.8). If there is no available space within the urban footprint the only options for developers are to expand the footprint, leading to urban sprawl, or to build higher and deeper. Brownfields provide essential flexibility within urban systems. As St Pancras welcomed its first Eurostars, the next use for Waterloo International was to be a return to domestic rail following a refurbishment due to take until 2014 (Glaspool 2009). The City of London has a high percentage, of an admittedly very small area, declared as PDL in its National Land Use Database returns, but that most of such "Brownfields" is A-type and rapidly redeveloped by private sector developers.

---

[2] Waterloo was the London terminus for the Eurostar between 14th November 1994 13th November 2007 when Paddington International opened.

Fig. 25.8 CABERNET Land use model: Brownfields are the empty space that allows urban systems to develop (source: CABERNET 2006, reproduced with permission)

Long recognised as good practice, encouraged by the Aarhus Convention (Anon 1998) and now a legal requirement of the Public Participation Directive (European Parliament 2003), citizen engagement is an essential component of ensuring the long term success of a master plan.

"Design charrettes" are one highly effective way of engaging with a wide range of citizens in a short period of time. By compressing the time allowed for ideas to emerge, participants focus attention on the big issues rather than the often irrelevant and surprisingly contentious details. Collaborative working in an environment of mutual respect and tolerance of ideas and suggestions leads to consensus or identification of fundamental areas of discord in a short time (Fig. 25.9).

Fig. 25.9 Recording the outcomes of a design charrette (source: C P Nathanail, reproduced with permission)

25.12 Applying a Systems Analysis Approach to Brownfield Redevelopment

A system is a group of natural or artificial things that connect to form a whole (Oxford English Dictionary 1989). Cities, towns, villages have been considered to be complex systems with interactions within and among their environmental, social and economic spheres (Nathanail 2005). Any form of management presumes understanding: a systems analysis approach is a useful way to gain such understanding (Alfeld and Meadows 1974; Nathanail 2005) (Leney AD (2008) A systems approach to assess the redevelopment options for urban Brownfield sites. PhD Thesis, University of Nottingham, Nottingham, “unpublished”). Brownfield redevelopment constitutes a system perturbation: the urban system will interact with and be altered by the redeveloped site. By considering the impact of different redevelopment options it is possible to assess how successfully each option meets the project objectives, and therefore which of the considered options is most appropriate for a site.

Nathanail (2005) and Leney (Leney AD (2008) A systems approach to assess the redevelopment options for urban Brownfield sites. PhD Thesis, University of Nottingham, Nottingham, “unpublished”) have demonstrated that the REMIT/RESPONSE methodology developed by Hudson (1992) and applied by Nathanail et al. (1992) to open cast coal mining can objectively analyse and help evaluate how perturbations such as new developments affect an urban system.

25.13 Opportunities for Synergy (e.g. Carbon, Energy and Waste Management)

Reclamation and remediation offer opportunities for energy efficient forms of construction to be incorporated in to the redevelopment. For example pump and treat systems can be used as heat exchangers. Excavations for ex situ treatment or off site disposal can be exploited to install ground source heat pump infrastructure or create extra space in the form of basements. Electric powered remediation can reduce the lifetime costs of installing renewable energy generation such as wind turbines or photovoltaic cells.

The Brownfield process manager should be adept at identifying and seeking out such synergy in order to maximise the benefit and minimise the cost of redevelopment.

25.14 Future Perspectives

This chapter is being finalised in the midst of the deep financial recession and in the face of rapidly rising oil prices. At present, structural change is happening faster in Western Europe and North America than for a very long time. Financial markets remain in turmoil, interest rates are at an historic low and the development industry going through one of its deepest troughs.

Such change creates Brownfields as car factories shut down, shops close and leisure facilities fail due to lack of visitors. (Former) giants such as AIG, General Motors and the Royal Bank of Scotland are in public ownership. High street names such as Woolworths have failed resulting in many vacant city centre stores.

The space for manoeuvre in the CABERNET land use puzzle has perhaps more flexibility than it truly needs and we would like it to have. After all the pieces of the puzzle need each other to avoid falling out. A system needs its component parts to continue to function.

Some argue that peak oil and climate change are tipping points which will result in a paradigm shift in the way urban societies manage themselves and their relationship with the rest of the planet. The Transition concept represents one bottom-up approach to ensuring local urban communities have the social and intellectual capital to survive energy descent. Brownfields are an essential local resource that such initiatives will have to plan to exploit wisely if they are to succeed by surviving.

One final thought is that US and potentially international approaches to contaminated site management arose from the response to the federal emergency that developing the former Love Canal gave rise to (Box 25.3).

Box 25.3 Failed Vision Creates Brownfields

William Love's vision of a Direct Current (DC) powered utopia on the shores of lakes Erie and Ontario failed. The advent of Alternating Current (AC) technology killed off the competing DC technology: Love' utopian dream ended and the part filled canal intended to carry water through the turbines was abandoned. Eventually the Hooker Chemical company bought this uncontaminated Brownfield site and disposed of its industrial wastes in it. Post World War II, a growing population meant the local Board of Education wanted more land for schools and housing: they bought land from Hooker and, ignoring institutional controls prohibiting development, built homes and a school on Love's waste infilled canal. This closure of the contaminated-pathway-receptor linkage came to national attention in the 1970s and resulted in the US Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) – more commonly know as the Superfund. Today Love Canal is synonymous with contaminated site management but it started out as an uncontaminated site that met the CABERNET definition of Brownfield well before it became a chemical waste repository.

References

Anon (1998) Convention on access to information, public participation in decision-making and access to justice in environmental matters, done at Aarhus, Denmark, on 25 June 1998 http://www.unece.org/env/pp/documents/cep43e.pdf. Accessed 29 June 2009

Alfeld LE, Meadows DL (1974) A systems approach to urban revival. In: Mass NJ (ed) Readings in Urban Dynamics, vol. 1. Wright–Allan Press, Cambridge, MA

Alker S, Joy V, Roberts P, Smith N (2000) The definition of brownfield. J Environ Plan Manag 43:49–69

CABERNET (2005) CABERNET position paper – professional skills in Brownfield regeneration, 1st edn – April 2005. Available from www.cabernet.org.uk. Accessed 24 August 2010

CABERNET (2006) Sustainable Brownfield regeneration: CABERNET network report. Millar K, Grimski D, Ferber U, Nathanail CP, University of Nottingham Land quality management report. ISBN 0-9547474-5-3 http://www.cabernet.org.uk/resourcefs/427.pdf. Accessed 29 June 2009

European Parliament (2003) Directive 2003/35/EC providing for public participation in respect of the drawing up of certain plans and programmes relating to the environment. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:156:0017:0024:EN:PDF. Accessed 29 June 2009

Ferber U, Grimski D (2002) Brownfields and redevelopment of urban areas. Austrian Federal Environment Agency on behalf of CLARINET, Vienna, Austria pp 9–10

Ferguson C, Darmendrail D, Freier K, Jensen BK, Jensen J, Kasamas H, Urzelai A, Vegter J (1998) Risk assessment for contaminated sites in Europe Vol 1: scientific basis. Land Quality, Nottingham ISBN 0 9533090 0 2. http://www.umweltbundesamt.at/fileadmin/site/umweltthemen/altlasten/caracas_publ1.pdf. Accessed 2 July 2009

Ferguson C, Kasamas H (1999) Risk assessment for contaminated sites in Europe Vol 2: policy frameworks. Land Quality, Nottingham. ISBN 0 9533090 1 0

Foxell S (2008) Mapping London: making sense of the city. Black Dog, London

Glancey J (2003) It's a jumble out there. The guardian Monday 3 November 2003. http://www.guardian.co.uk/artanddesign/2003/nov/03/architecture.regeneration. Accessed 2 July 2009

Glaspool D (2009) Waterloo international. http://www.kentrail.co.uk/waterloo_international.htm. Accessed 29 June 2009

Hansard (2008) Gardens. http://www.publications.parliament.uk/pa/cm200809/cmhansrd/cm090402/text/90402w0048.htm#0904032000060. Accessed 29 June 2009

Hudson JA (1992) Rock engineering systems: theory and practice, Ellis Horwood, Chichester

Kok W (2004) Facing the challenge: the Lisbon strategy for growth and employment. Report from the High Level Group chaired by Wim Kok. http://europa.eu.int/comm/lisbon_strategy/index_en.html. Accessed 24 August 2010

Land Restoration Trust (2009) Our approach (restoring land for the community). http://www.landrestorationtrust.org.uk/template.asp?l1=2&l2=87. Accessed 29 June 2009

Nathanail CP (2005) A systematic approach to delivering sustainable urban brownfield regeneration. CABERNET 2005: the international conference on managing urban land, Belfast, Land Quality, Nottingham

Nathanail CP, Bardos RP (2004) Reclamation of contaminated land. Wiley, Chichester

Nathanail CP, Earle DA, Hudson JA (1992). Stability hazard indicator system for slope failure in heterogeneous strata. In: EUROCK '92 symposium, Chichester, UK, Thomas Telford, pp 111–116

National Trust (2009) Our past. http://www.nationaltrust.org.uk/main/w-trust/w-thecharity/w-thecharity_our-past.htm. Accessed 29 June 2009

Oxford English Dictionary (1989) 2nd ed. Clarendon Press, Oxford

Regen WM (2009) http://www.regenwm.org/. Accessed 29 June 2009

RESCUE (2005) Best practice guidance for sustainable Brownfield regeneration. Edwards D, Pahlen G, Bertram C, Nathanail C P, Land Quality, Nottingham. http://www.rescue-europe.com/download/reports/RESCUE%20Manual.pdf. Accessed 29 June 2009

Rogers R (2007) London as it could be. http://www.richardrogers.co.uk/render.aspx?siteID=1&navIDs=1,4,22,562. Accessed 29 June 2009

Umweltbundesamt (2009) CARACAS – concerted action for risk assessment for contaminated sites in Europe http://www.umweltbundesamt.at/en/umweltschutz/altlasten/projekte1/international1/caracas. Accessed 2 July 2009

Index

F.A. Swartjes (ed.), *Dealing with Contaminated Sites*,
DOI 10.1007/978-90-481-9757-6,