

Emilio Chuvieco
Jonathan Li
Xiaojun Yang
(Eds.)



Advances in

Earth Observation

Advances in Earth Observation of Global Change

Emilio Chuvieco · Jonathan Li · Xiaojun Yang
Editors

Advances in Earth Observation of Global Change

 Springer

Editors

Prof. Dr. Emilio Chuvieco
Universidad Alcalá de Henares
Department of Geography
Depto. Geografía
Calle Colegios 2
22801 Alcalá de Henares
Spain
emilio.chuvieco@uah.es

Prof. Dr. Jonathan Li
University of Waterloo
Department of Geography
and Environmental Management
200 University Avenue West
Waterloo, Ontario N2L 3G1
Canada
junli@uwaterloo.ca

Prof. Xiaojun Yang
Florida State University
Dept. Geography
Tallahassee Florida 32306
USA
xyang@fsu.edu

ISBN 978-90-481-9084-3 e-ISBN 978-90-481-9085-0

DOI 10.1007/978-90-481-9085-0

Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010932447

© Springer Science+Business Media B.V. 2010

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

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

Preface

In the last years, public opinion and political managers increasingly consider global change studies as a vital field to better address the current environmental challenges. Climate alterations, natural hazards, biodiversity, water resources and other critical phenomena need to be understood at different spatial scales, from local to global. Reliable and long-term databases are indispensable in these studies, to establish temporal and spatial trends that will provide a better conception of the sources and impacts of changes.

Earth Observation of global changes is increasingly considerable as an indispensable tool to have such a comprehensive view of current problems and challenges. Images acquired by EO satellites provide a synoptic and repetitive inspection on the state of natural resources. Sensor systems currently acquired images from different spectral bands (visible, near, medium and thermal infrared and microwave), with a wide range of spatial resolutions (from 1 m to 3 km pixel size), and at different temporal intervals (from 15 min to several weeks). This information may be easily integrated into geographical databases, thus providing an ideal scenario for spatial planning at different spatial scales.

This volume gathers a selection of papers presented at the Second International Conference on Earth Observation of Global Change (EOGC2009), held in Chengdu, China, 26–29 May, 2009. The first conference was held in Madrid, Spain, 2 years earlier and was also published by Springer (Chuvieco E (ed) (2008) *Earth Observation of Global Change. The role of satellite remote sensing in monitoring the Global Environment*. Springer). This present volume is organized around six topics:

1. Remote Sensing of Land Use and Land Cover Change ([Chapters 1–4](#)).
2. Remote Sensing of Coastal and Marine Ecosystems ([Chapters 5 and 6](#)).
3. Remote Sensing of Snow, Ice, and the Polar Environment ([Chapters 7–9](#)).
4. Observing Global Change by Geodetic Techniques ([Chapters 11–12](#)).
5. Earth Observation for Natural Hazards Monitoring and Assessment ([Chapters 13–16](#)).
6. Geospatial Data Processing and Integration for Change Detection ([Chapters 17–19](#)).

All papers have been reviewed by external scientists and prepared specifically to be part of this book. The editors in charge of this process have been Emilio Chuvieco (Sections 1 and 5), Jonathan Li (Sections 4 and 6) and Xiaojun Yang (Sections 2 and 3).

The final selection of papers includes a wide range of topics, where Earth Observation data have proven to be very valuable for analysis and planning: urban growth, natural hazards (earthquakes, landslides, forest fires, etc.), land degradation process, water quality, ice and snow monitoring, tectonic motion, ocean height, food security, and weather patterns. These variables have been estimated with a great variety of sensors: from optical to thermal and microwave, and satellite missions: Terra-MODIS, Envisat-MERIS, Landsat-TM/ETM+, ALOS-SAR, Envisat-SAR, Radar Altimeters, among others. This variety of data will be extended in the near future with many other missions that are being planned by the different space agencies. It is worth to mention the European Space Agency's Sentinel program, as well as the continuation of NASA Mission to Planet Earth. Meteorological agencies will also provide further relevant information from the geostationary and polar orbiting satellites. This panorama will increase in the near future with the missions coming from the private sector, which are being benefited from the reduction in launching costs, optics and electronic components, providing an even wider range of on-demand sensors based and small satellites.

This whole flow of information needs to be properly ingested in current climate and vegetation models, to reduce the uncertainty of modelling future scenarios of change. The increase use of EO data will certainly make model outputs more versatile, precise and accurate. Therefore, EO data will continue to be a critical source of spatial information to better understand and preserve our fragile Planet.

Madrid, Spain
Waterloo, Ontario, Canada
Tallahassee, Florida

Emilio Chuvieco
Jonathan Li
Xiaojun Yang

Contents

1	The Role of Small Satellite Missions in Global Change Studies . . .	1
	Rainer Sandau and Klaus Briess	
2	Spatial Pattern Analysis of Water-Driven Land Cover Change in Aridzone, Northwest of China	17
	Qiming Zhou and Bo Sun	
3	A Spatial Logistic Regression Model for Simulating Land Use Patterns: A Case Study of the Shiraz Metropolitan Area of Iran	27
	Amin Tayyebi, Mahmoud Reza Delavar, Mohammad Javad Yazdanpanah, Bryan Christopher Pijanowski, Sara Saeedi, and Amir Hossein Tayyebi	
4	Abu Dhabi Island: Analysis of Development and Vegetation Change Using Remote Sensing (1972–2000)	43
	Abdulkhakim M. Abdi and Anand Nandipati	
5	Remote Sensing of Coastal Water Quality in the Baltic Sea Using MERIS	55
	Andreas Neumann, Harald Krawczyk, and Stefan Riha	
6	Remote Sensing for Water Quality Monitoring in Apalachicola Bay, USA	69
	Wenrui Huang, Shuisen Chen, and Xiaojun Yang	
7	Extracting Cryospheric Information over Lowlands from L-Band Polarimetric SAR Data	79
	Nicolas Longép�, Masanobu Shimada, Sophie Allain, and Eric Pottier	
8	Variability of Northern Hemisphere Spring Snowmelt Dates Using the AVHRR Polar Pathfinder Snow Cover During 1982–2004	97
	Hongxu Zhao and Richard Fernandes	

9 MODIS Snow Monitoring Over the Tibetan Plateau 111
 Li Xu

**10 The Global Geodetic Observing System (GGOS):
 Detecting the Fingerprints of Global Change
 in Geodetic Quantities 125**
 Hans-Peter Plag, Chris Rizos, Markus Rothacher, and
 Ruth Neilan

**11 Monitoring Radial Tectonic Motions of Continental
 Borders Around the Atlantic Ocean and Regional Sea Level
 Changes by Space Geodetic Observations 145**
 Zhigen Yang and Fengchun Shu

**12 GNSS Activities for Natural Disaster Monitoring and
 Climate Change Detection at GFZ – An Overview 159**
 Junping Chen, Michael Bender, Geory Beyerle, Galina Dick,
 Carsten Falck, Maorong Ge, Gerd Gendt, Stefen Heise,
 Marbus Ramatschi, Torsten Schmidt, Ralf Stosius, and
 Jens Wickert

**13 Satellite Imagery for Landslide Mapping in an Earthquake-
 Struck Area 173**
 Xiaojun Yang

14 Relations Between Human Factors and Global Fire Activity 187
 Emilio Chuvieco and Chris Justice

**15 The Use of Remote Sensing Data and Meteorological
 Information for Food Security Monitoring, Examples in
 East Africa 201**
 Michel Massart, Felix Rembold, Oscar Rojas, and Olivier Leo

16 Application of an Early Warning System for Floods 217
 Adriana Albanese, Piero Boccoardo, Fabio Giorgi,
 Nishshanka Prasanna Premachandra, Olivier Terzo, and
 Rossella Vigna

**17 L-Band and C-Band Combined Interferometric Monitoring
 of the Wenchuan Earthquake 239**
 Kui Zhang, Alex Hay-Man Ng, Linlin Ge, Yusen Dong, and
 Chris Rizos

**18 Uncovering the Space–Time Patterns of Change with the
 Use of Change Analyst – Case Study of Hong Kong 255**
 Bo Huang and Hua-Leung Sin

**19 Change Detection of Sea Ice Distribution in SAR Imagery
 Using Semi-variogram of Intrinsic Regionalization Model 269**
 Yu Li and Jonathan Li

Index 281

Contributors

Abdulkhikim M. Abdi Instituto Superior de Estatística e Gestão de Informação, Universidade Nova de Lisboa, Campus de Campolide, 1070-312 Lisboa, Portugal, hakim.abdi@gmail.com

Adriana Albanese DITAG, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy, adriana.albanese@polito.it

Sophie Allain Institute of Electronics and Telecommunications of Rennes, University of Rennes I, Bât. 11D, 263 av. Général Leclerc, 35042 Rennes, France, sophie.allain@univ-rennes1.fr

Michael Bender Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Geary Beyerle Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Piero Boccardo ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action), Via P.C. Boggio 61, Torino, Italy, piero.boccardo@polito.it

Klaus Briess Institut für luft- und Raumfahrt, TU Berlin, Marchstrasse 12/Sekr. Fb.10587 Berlin, Germany, klaus.briess@ilr.tu-berlin

Junping Chen Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany, junping.chen@gfz-potsdam.de

Shuisen Chen Guangzhou Institute of Geography, Guangzhou, Guangdong 510070, China, css@gdas.ac.cn

Emilio Chuvieco Department of Geography, University of Alcalá, Colegios 2, 28801 Alcalá de Henares, Spain, emilio.chuvieco@uah.es

Mahmoud Reza Delavar Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran

Galina Dick Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Yusen Dong Faculty of Earth Sciences, China University of Geosciences, Wuhan, Hubei 430074, China

Carsten Falck Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Richard Fernandes Earth Sciences Sector, Canada Centre for Remote Sensing, Natural Resources Canada, Ottawa, Canada, Richard.Fernandes@ccrs.nrcan.gc.ca

Linlin Ge School of Surveying and Spatial Information Systems, The University of New South Wales, Kensington, Sydney, NSW 2025, Australia, l.ge@unsw.edu.au

Maorong Ge Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Gerd Gendt Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Fabio Giorgi ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action), Via P.C. Boggio 61, Torino, Italy, fabio.giorgi@ithaca.polito.it

Stefan Heise Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Bo Huang Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China, bohuang@cuhk.edu.hk

Wenrui Huang Department of Civil Engineering, Florida State University, Tallahassee, FL 32310, USA, whuang@eng.fsu.edu

Chris Justice Geography Department, University of Maryland, College Park, MD 20742, USA, justice@hermes.geog.umd.edu

Harald Krawczyk Remote Sensing Technology Institute, German Aerospace Center DLR, Rutherfordstrasse 2, 2489 Berlin, Germany, harald.krawczyk@dlr.de

Olivier Leo EU-DG JRC, Institution for the Protection and Security of the Citizen, MARS Unit TP 266, Via E. Fermi, 21020 Ispra (VA), Italy, Olivier.leo@jrc.ec.europa.eu

Jonathan Li Department of Geography and Environmental Management, Faculty of Environment, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada, junli@uwaterloo.ca

Yu Li Department of Geography and Environmental Management, Faculty of Environment, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada, y62li@uwaterloo.ca

Nicolas Longépé Earth Observation Research Center, Japan Aerospace Exploration Agency, Sengen 2-1-1, 305-8505 Tsukuba, Ibaraki, Japan, nicolas.longepe@jaxa.jp

Michel Massart EU-DG JRC, Institution for the Protection and Security of the Citizen, MARS Unit TP 266, Via E. Fermi, 21020 Ispra (VA), Italy, michel.massart@jrc.ec.europa.eu

Anand Nandipati Instituto Superior de Estatística e Gestão de Informação, Universidade Nova de Lisboa, Campus de Campolide, 1070-312 Lisboa, Portugal, anand6@gmail.com

Ruth Neilan IGS Central Bureau, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

Andreas Neumann Remote Sensing Technology Institute, German Aerospace Center DLR, Rutherfordstrasse 2, 2489 Berlin, Germany, andreas.neumann@dlr.de

Alex Hay-Man Ng Cooperative Research Centre for Spatial Information and School of Surveying and Spatial Information Systems, The University of New South Wales, Kensington, Sydney, NSW 2052, Australia, z3021790@student.unsw.edu.au

Bryan Christopher Pijanowski Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana, USA

Hans-Peter Plag Nevada Bureau of Mines and Geology and Seismological Laboratory, University of Nevada at Reno, Mail Stop 175, Reno, NV 89557, USA, hpplag@unr.edu

Eric Pottier Institute of Electronics and Telecommunications of Rennes, University of Rennes I, Bât. 11D, 263 av. Général Leclerc, 35042 Rennes, France, eric.pottier@univ-rennes1.fr

Nishshanka Prasanna Premachandra DITAG, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy, prasanna.nishshanka@polito.it

Markus Ramatschi Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Felix Rembold EU-DG JRC, Institution for the Protection and Security of the Citizen, MARS Unit TP 266, Via E. Fermi, 21020 Ispra (VA), Italy, Felix.rembold@jrc.ec.europa.eu

Stefan Riha Remote Sensing Technology Institute, German Aerospace Center DLR, Rutherfordstrasse 2, 2489 Berlin, Germany, stefan.riha@dlr.de

Chris Rizos School of Surveying and Spatial Information Systems, The University of New South Wales, Kensington, Sydney, NSW 2052, Australia, c.rizos@unsw.edu.au

Oscar Rojas EU-DG JRC, Institution for the Protection and Security of the Citizen, MARS Unit TP 266, Via E. Fermi, 21020 Ispra (VA), Italy, Oscar.rojas@jrc.ec.europa.eu

Markus Rothacher Institute of Geodesy and Photogrammetry, ETH Zürich, Schafmattstr. 34/HPV G52, CH 8093 Zürich, Switzerland, markusrothacher@ethz.ch

Sara Saeedi Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Iran

Rainer Sandau DLR, German Aerospace Center, Rutherfordstr. 2, 12489 Berlin, Germany, rainer.sandau@dlr.de

Torsten Schmidt Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Masanobu Shimada Earth Observation Research Center, Japan Aerospace Exploration Agency, Sengen 2-1-1, 305-8505 Tsukuba, Ibaraki, Japan, shimada.masanobu@jaxa.jp

Fengchun Shu Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China

Hua-Leung Sin Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China, william.sin@cuhk.edu.hk

Ralf Stosius Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Bo Sun Department of Geography, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong S.A.R, China

Amin Tayyebi Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran, amin.tayyebi@gmail.com

Amir Hossein Tayyebi Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran

Olivier Terzo ISMB (Istituto Superiore Mario Boella), Via P.C. Boggio 61, Torino, Italy, terzo@ismb.it

Rosella Vigna DITAG, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy, rossella.vigna@polito.it

Jens Wickert Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany

Li Xu National Climate Center, China Meteorological Administration, Beijing 100081, China; Department of Atmospheric, Oceanic and Earth Science, George Mason University, Fairfax, VA 22030, USA, lxu2@gmu.edu; Institute of Global Environment and Society, Calverton, MD 20705, USA

Xiaojun Yang Department of Geography, Florida State University, Tallahassee, FL 32306, USA, xyang@fsu.edu

Zhigen Yang Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China, yangz@shao.ac.cn

Mohammad Javad Yazdanpanah School of Electrical and Computer Engineering, Control and Intelligent Processing Center of Excellence, College of Engineering, University of Tehran, Tehran, Iran

Kui Zhang Cooperative Research Centre for Spatial Information and School of Surveying and Spatial Information Systems, The University of New South Wales, Kensington, Sydney, NSW 2052, Australia, kubic.zhang@gmail.com

Hongxu Zhao Adaptation and Impacts Research Section, Climate Research Division, Environment Canada, Toronto, Canada, hongxu.zhao@ec.gc.ca

Qiming Zhou Department of Geography, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong S.A.R, China, qiming@hkbu.edu.hk

Chapter 1

The Role of Small Satellite Missions in Global Change Studies

Rainer Sandau and Klaus Briess

Abstract There is an increasing need for Earth Observation (EO) missions to meet the information requirements in connection with Global Change Studies. Small and cost-effective missions are powerful tools to flexibly react on information requirements with space borne solutions. Small satellite missions can be conducted relatively quickly and inexpensively and provide increased opportunity for access to space. The spacecraft bus and instruments can be based either on optimised off-the-shelf systems, with little or no requirements for new technology, or on new high-technology systems. Thus a new class of advanced small satellites, including autonomously operating “intelligent” satellites may be created, opening new fields of application for scientific purposes as well as operational, public and commercial services. Further milestones in the small satellite Earth observation mission developments are the availability and improvement of small launchers, the development of small ground station networks connected with rapid and cost-effective data distribution methods, and cost-effective management and quality assurance procedures. Advantages of small satellite missions, complementing the large complex missions are: more frequent mission opportunities and therefore faster return of science and application data; larger variety of missions and therefore also greater diversification of potential users; more rapid expansion of the technical and/or scientific knowledge base; and greater involvement of local and small industry. The paper deals with general trends in the field of small satellite missions for Earth observation. Special attention is given to the potential of spatial, spectral, and temporal resolution of small satellite based systems. Examples show that constellations give the unique for small satellites possibility to provide good daily coverage of the globe or/and allow to observe dynamic phenomena.

R. Sandau (✉)

DLR, German Aerospace Center, Rutherfordstr. 2, 12489 Berlin, Germany
e-mail: rainer.sandau@dlr.de

1.1 Introduction

The high need for Earth observation missions in order to improve the data base contents in connection with global change studies is perhaps most clearly seen in the many current moves for international co-operation in the field of environment where measurements from Earth observing satellites are an essential element. This is especially so where we need to acquire, analyze and use data documenting the condition of the Earth's resources and environment on a long-term (permanent) basis.

In 2008 the Group on Earth Observations, which currently numbers some 74 countries, the European Commission and 51 participating organizations, has concrete plans for its Global Earth Observation System of Systems (GEO, 2008). In 2008 the European Union's Space Council continued to advance Europe's Space Policy, reaffirming the need for rapid implementation of the Global Monitoring for Environment and Security (GMES) (GMES, 2009a, b) programme.

From the space-borne remote sensing point of view, the only way to meet the information requirements is to pursue activities to develop and operate cost-effective Earth observations missions, especially small satellite missions.

Small satellite missions can be achieved by using different approaches and methods. One approach is to focus on a single task and use available off-the-shelf technology to build a small satellite system (bus and payload) for the intended remote sensing purpose. Another possible approach is to take full advantage of the ongoing technology developments leading to further miniaturization of engineering components, development of micro-technologies for sensors and instruments which allow designing dedicated, well-focused high-performance Earth observation missions.

Since the advent of modern technologies, small satellites using off-the-shelf technologies or missions focused on specific physical phenomena have also been perceived to offer an opportunity for countries with a modest research budget and little or no experience in space technology, to enter the field of space-borne Earth observation and its applications. Small satellite technology is a major mean to bring within the reach of every country the opportunity to operate small satellite Earth observation missions and utilize the data effectively at low costs, as well as to develop and build application-driven missions. It provides the opportunity to conduct or participate in Earth observation missions using small, economical satellites, and associated launches, ground stations, data distributions structures, and space system management approaches.

The situation in the field of small satellite missions for Earth observation has matured in the last 10 years. This may be, for instance, observed from the topics and the quality of contributions to the series of conferences taking place in Berlin, Logan, at the annual International Astronautical Congress or conducted by space agencies like ESA or CNES.

But what exactly is a small satellite? The International Academy of Astronautics IAA proposed a simplified definition (Sandau, 2006). This definition is reflected in Fig. 1.1 in conjunction with additional features which are essential when discussing

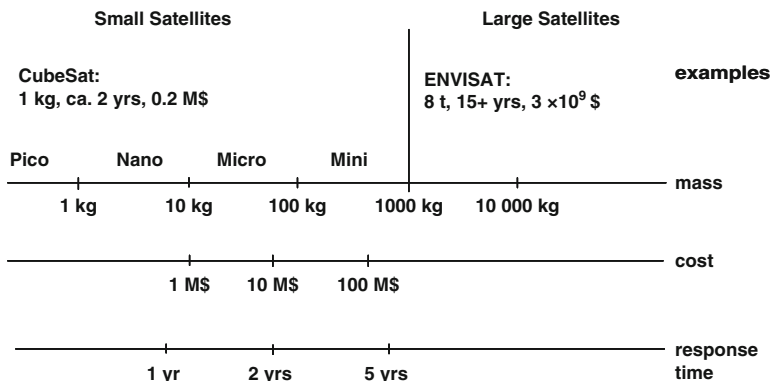


Fig. 1.1 Some features of small satellites

small satellite characteristics like cost and response time. The performance issue is covered in the subsequent chapters of the paper.

At UNISPACE III (Background Paper, 1998), the costs of developing and manufacturing a typical mini-satellite was indicated to be US\$ 5–20 million, while the cost of a micro-satellite was correspondingly US\$ 2–5 million. The cost of a nano-satellite could be below US\$ 1 million (prices as of 1999). Whereas the development and production time for large satellites is observed to be 15+ years, the corresponding time for minis should be 3–5 years, for micros 1.5 years, for nanos about 1 year, and for picos less than 1 year. Of course, cost and duration figures are to be considered ball park figures. They are bases on the usage of state-of-the-art technology by professional teams. They may deviate considerably if key technology is to be developed and/or the implementation teams are at the beginning of the learning curve. Figure 1.1 is complemented by two examples showing the edges of the feature ranges.

In the next chapters of the paper, the actual status of small satellite missions is summarized. Small satellites of Surrey Satellite Technology Ltd. SSTL, the small satellite initiatives of the USA and in the rest of the world are addressed, as well as small satellites developed by universities. Constellations are another aspect considered.

1.2 Small Satellite Missions: Facts and Trends

From Fig. 1.1 we can learn that, roughly, the smaller the satellite the less the cost and the response time. This is a strong motivation to try to go for small satellite missions. The International Academy of Astronautics IAA study (Sandau, 2006) presented the state of the art of small satellite missions and examined more factors that enable one to produce a cost-effective small satellite mission for Earth observation. It seems,

while there are several examples of such missions flying today, the lessons that must be learned in order to produce cost-effective small sat missions have neither been universally accepted nor understood by all in the space community. One of the intentions of that study was to point out how a potential user can produce a cost effective mission. One of the key enablers of designing a cost-effective mission is having the key expertise available. As the number of successfully space-faring nations grows, the pool of expertise available to meet the challenges of small mission grows.

1.2.1 General Facts

Small satellite missions can be achieved by using different approaches and methods.

Since about two decades, small satellites have also been perceived to offer an opportunity for countries with a modest research budget and little or no experience in space technology, to enter the field of space-borne Earth observation and its applications. Small satellite technology is a major mean to bring within the reach of every country the opportunity to operate small satellite Earth observation missions and utilize the data effectively at low costs, as well as to develop and build application-driven missions. It provides the opportunity to conduct or participate in Earth observation missions using small, economical satellites, and associated launches, ground stations, data distributions structures, and space system management approaches.

One of the possible approaches is taking full advantage of the ongoing technology developments leading to further miniaturization of engineering components, development of micro-technologies for sensors and instruments which allow to design dedicated, well-focused Earth observation missions. At the extreme end of the miniaturization, the integration of micro-electromechanical systems (MEMS) with microelectronics for data processing, signal conditioning, power conditioning, and communications leads to the concept of application specific integrated micro-instruments (ASIM). These micro- and nano-technologies have led to the concepts of nano- and pico-satellites, constructed by stacking wafer-scale ASIMs together with solar cells and antennas on the exterior surface, enabling the concept of space sensor webs.

The advantages of small satellite missions are:

- more frequent mission opportunities and therefore faster return of science and for application data
- larger variety of missions and therefore also greater diversification of potential users
- more rapid expansion of the technical and/or scientific knowledge base
- greater involvement of local and small industry.

After some years of global experience in developing low cost or cost-effective Earth observation missions, one may break down the missions into categories like:

- Commercial – Requiring a profit to be made from satellite data or services
- Scientific/Military – Requiring new scientific/military data to be obtained
- New technology – Developing or demonstrating a new level of technology
- Competency demonstration – Developing and demonstrating a space systems competency
- Space technology transfer/training – Space conversion of already competent engineering teams
- Engineering competency growth – Developing engineering competence using space as a motivation
- Education – Personal growth of students via course projects or team project participation

Large satellite missions and small satellite missions are considered to be complementary rather than competitive. The large satellite missions are sometimes even a precondition for cost-effective approaches.

1.2.2 Trends

Small satellite missions are supported by several contemporary trends:

- Advances in electronic miniaturization and associated performance capability;
- The recent appearance on the market of new small launchers (e.g. through the use of modified military missiles to launch small satellites);
- The possibility of “independence” in space (small satellites can provide an affordable way for many countries to achieve Earth Observation and/or defense capability, without relying on inputs from the major space-faring nations);
- Ongoing reduction in mission complexity as well as in those costs associated with management; with meeting safety regulations etc.;
- The development of small ground station networks connected with rapid and cost-effective data distribution methods.

1.3 Resolution Requirements for Space Borne Remote Sensing

When we discuss the performance parameters, we address in the first hand the resolution in terms of:

- Spatial resolution
- Spectral resolution
- Temporal resolution

For space borne sensors, further requirement may be considered concerning mass, volume and power consumption. These are essential features when it comes to the spacecraft design, especially when targeting to small satellites. Figures 1.2 and 1.3 show the very divers requirements connected with the different remote

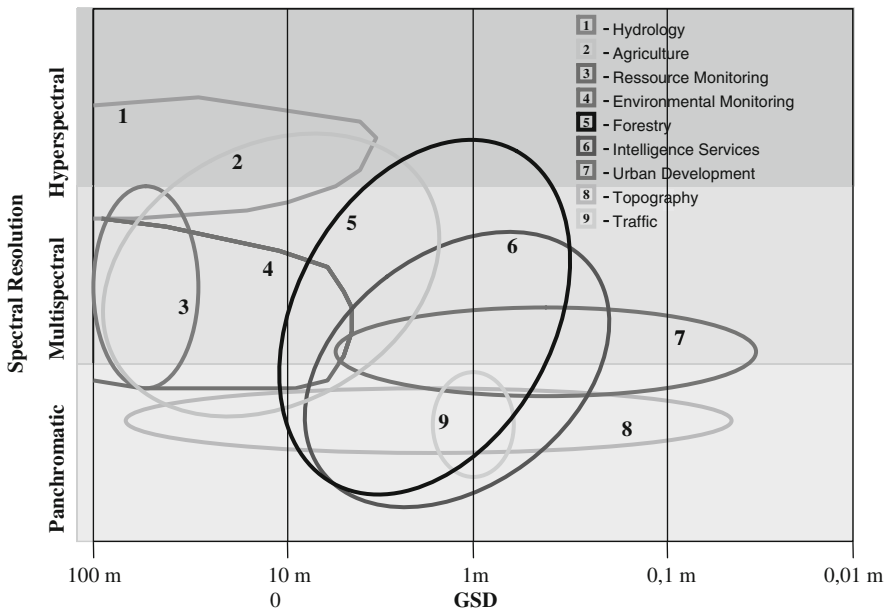


Fig. 1.2 Earth observation request: ground sampling distance GSD versus spectral resolution

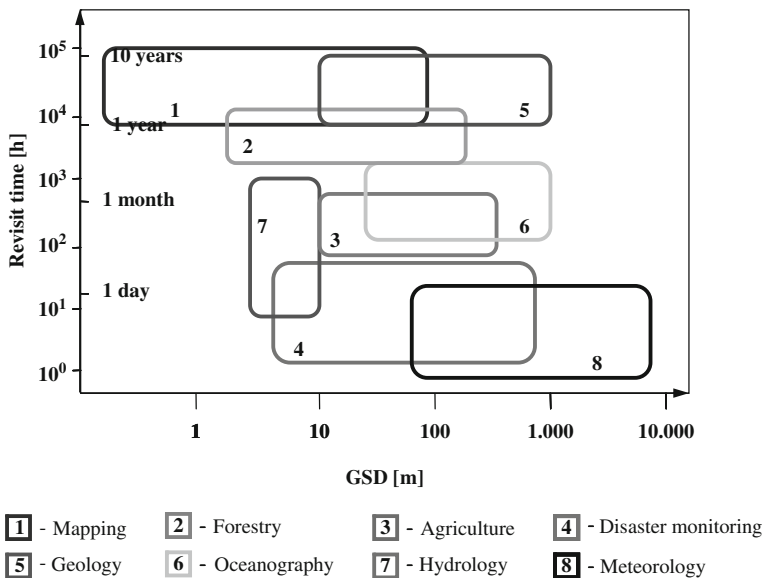


Fig. 1.3 Earth observation request: GSD vs revisit time

sensing application fields (Roeser et al., 2000). The range of spatial resolution, the ground sampling distance GSD, covers centimeters to several hundred meters. The revisit time (temporal resolution) ranges from less than 1 h to 10 years. The range of spectral resolution requirements starts with panchromatic only for topographic mapping and ends with hyper-spectral resolution, for instance in the field of hydrology.

Figure 1.2 relates the spectral requirements to the ground sampling distance GSD for different application fields. It can be seen that some important applications like agriculture, environmental monitoring, resource monitoring, topography and hydrology require a moderate ground pixel size in the order of 10–100 m but several spectral channels (multispectral or hyper-spectral instruments). These requirements can be met by multi-spectral cameras, imaging spectrometer or imaging spectro-radiometer fitting to small satellites.

Figure 1.3 depicts temporal resolution versus ground sampling distance GSD for different application fields. Only few applications require a high repeating time of one or few days like disaster monitoring. To meet the requirement of a high revisit time with optical instruments with a limited swath width because of the required moderate or high ground sampling distance only a number of satellites flying in a constellation are the solution. These satellites are preferably small satellites because of the affordability. Especially data from disaster monitoring satellites are interesting for long term studies of disaster occurrence in time and location that means also for global change studies.

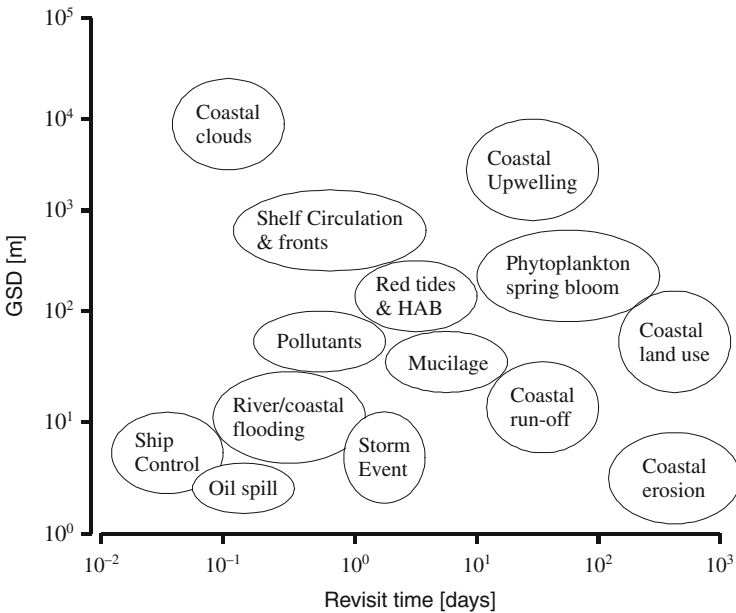


Fig. 1.4 Spatial and temporal requirements for coastal studies (after Hoepffner)

Even within the different application fields, the related subtasks may cover huge areas again. Figure 1.4 shows the requirement range for different coastal applications. In Fig. 1.4 the GSD ranges from meters to kilometers with revisit times from half an hour to several years (Sandau, 2006). Several coastal applications like coastal upwelling, coastal flooding, coastal run-off, storm event, coastal land use and phytoplankton spring bloom are monitored by small satellites. And these applications are of high importance also for global change studies because of increasing anthropogenic impact in these fields.

It is obvious: a small satellite system for remote sensing needs to focus on one of the application fields and within that application field on a specific task or group of tasks where feasible. With this background, when talking about a small satellite we are basically talking about an instrument or a complex instrument system optimized for a specific task and accommodated on a small satellite platform.

1.4 Resolution Capabilities of Small Satellite Systems

Seeing the huge diversity in the three different resolution areas spatial, spectral, and temporal, we need to answer two questions (for optical remote sensing systems) for small satellites and their payloads:

- what are the limitations, and
- what are the strengths

1.4.1 Spatial Resolution

The spatial resolution is increasing, i.e., the ground sampling distance GSD is decreasing. For example: the camera PIC-2 on the small satellite EROS-B from Israel provides a GSD of 0.70 m (Defense Update, 2006). EROS-B with a mass of 290 kg was already launched on 25.04.2006 with a Russian START-1 launcher into 500 km sun synchronous orbit (SSO).

1.4.2 Spectral Resolution

Also the spectral resolution is increasing. As an example may serve the hyper-spectral imager CHRIS on the ESA funded PROBA satellite (ESA, 2002). CHRIS, the 14 kg/9 W hyper-spectral imager, has a GSD of 18 m and provides up to 19 out of a total of 62 spectral bands in the VIS/NIR spectral range (400–1000 nm). PROBA with a mass below 100 kg (so it is a micro satellite) was launched into a 600 km SSO on 2.10.2001 together with the DLR/Germany micro satellite BIRD for forest fire detection and fire parameter assessment (Brieß et al., 2003), and the main payload TES (India) (EO, 2001) with the PSLV-C3 launcher from India.

1.4.3 Temporal Resolution

Small satellites provide a unique opportunity for affordable constellations. In this respect, small satellites can do things that are not practical with large satellites. At this point, the disaster monitoring constellation DMC may serve as the example for a constellation of five small satellites. We will elaborate more on constellations and formations in [Chapter 6](#) of this paper.

1.5 Status of Small Satellites

If we try to count the small satellites launched into space within the last 4 years, we achieve about the following results (for more details see Kramer, 2002):

- Company SSTL 14
- USA 15
- Rest of the World 60
- Universities 6, without the PicoSats.

All these small satellites are equipped with payloads to answer questions coming from different user groups. The favorite payloads for small satellites are cameras and spectrometers in the visible and near infrared wavelength range. According to this instrumentation most of the users groups are investigate the land surface to study albedo and reflectance characteristics, vegetation indices, land cover characteristics, urban regions, cartography and other multi-purpose imagery. But also the evaluation of ocean colors plays a role in determining sea surface parameters with multi-spectral cameras or spectrometers on small satellites. Another group of instruments especially suitable for small satellites are GNSS (Global Navigation Satellite System) receiver, electron detectors, magnetometers and other to measure parameters of atmosphere including the ionosphere, like the Total Electron Content, plasma density and other. Up to now the group active and passive radar instruments fit usually not to the resources of small satellites especially in the power consumption and geometric restrictions. There are some exceptions like the Indian RISAT satellite; it is a 300 kg RADAR satellite for reconnaissance. The status today is that no small RADAR satellite is available for Earth global change studies, small in the sense of a total satellite mass less than 500 kg.

Table 1.1 gives an overview to the status of small satellites suitable for global change studies. The DMC Beijing-1 has an optical payload and is part of the Disaster Monitoring Constellation. The EROS A, B and C are commercial small satellites for Earth observation (Israel).

HJ-1A and HJ-1B are small satellites from China dedicated to monitoring environment changes and the analysis and evaluation of natural disasters and environmental pollution. They have multi- and hyper-spectral instruments in the visible wavelength range on board an infrared camera system (HJ-1B). The HJ-1C will follow with a RADAR payload. The RapidEye constellation is described in the

Table 1.1 Planned in-orbit operations time of running and planned small satellite missions for Earth observation suitable for global change studies, TOPSAT – extended lifetime (different sources)

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
DMC BEIJING-1	[Bar]									
EROS A	[Bar]									
EROS B	[Bar]									
EROS C		[Bar]								
HJ-1A		[Bar]								
HJ-1B		[Bar]								
HJ-1C			[Bar]							
Rapid Eye		[Bar]								
TARANIS						[Bar]				
TOPSAT	[Bar]									
UK-DMC-2		[Bar]								

following section. The TARANIS satellite (CNES, France) will study the atmosphere with emphasis of the thermosphere. TOPSAT is a high resolution (2.5 m) optical satellite for Earth remote sensing from the United Kingdom. The UK-DMC-2 is the follow-on satellite from the United Kingdom for the next Disaster Monitoring Constellation with a multi-spectral camera.

Some typical mission parameters of operational small satellites (except VENU S in ca. 2012) suitable for global change studies are summarized in Table 1.2. Multi-spectral optical instruments including infrared are the most favorite payloads.

In future, as already partly in the past, small satellites may be equipped with payloads dedicated to answer specific questions coming from the teams dealing with the global change studies. In the future more small satellites for Earth remote sensing will fly in formation or constellation to assure a high area or time coverage or a high coverage of the same area by different instruments. Especially constellations, the unique domain of small satellites, give the chance to include also the dynamics of physical phenomena in the studies. For these reason two state-of-the-art small satellite constellations are described in the following chapter.

1.6 Constellations

Besides the potential high performance in terms of GSD and spectral resolution, constellations provide support also for observing dynamic characteristics of phenomena and/or a good daily coverage of the globe. Two examples are given here to show the potential: DMC-1 and successors, and RapidEye. These constellations are not specifically designed to support global change studies, but they may be used for such studies. At least they may be used as a model for future space borne systems dealing with global change aspects.

Table 1.2 Typical mission parameters of 5 small satellite missions suitable for global change studies

	TOPSAT	HJ-1A,-1B	Rapid Eye	UK-DMC-2	VENµS
Applications	Mapping, ad-hoc imagery, disaster monitoring (earthquake, land slides, other)	Monitoring environment changes, analysis and evaluation of natural disasters and environmental pollution	Orthographic Mapping, agricultural monitoring, Precision farming, Security, surveillance	Disaster Monitoring Constellation, Earth Observation	Land surface monitoring, To monitor influence of environmental factors (climate, topography, soils, humans)
Launch	27.10.2005	06.09.2008	29.08.2008	29.07.09	ca.2012
Number of satellites	1	2(+1 in 2010)	5	1(+3)	1
Av. Orbit height	689 km × 713 km	649 km	630 km	632 km × 685 km	720 km
Inclination	98.0°	97.9°	97.8°	98.1°	98.27°
Design life time	1 year	3 years	7 years	7 years	2.5(+1) years
Satellite mass	120 kg	470 kg	150 kg	120 kg	250 kg
Payload	MD camera	Hyper spectral & MS cameras, IR camera	MS camera VIS-NIR	MS camera	MS camera
Number of spectral bands	1 PAN, 3 MS	115 HS, 4 MS, 4 IR	5 MS	3 MS (VIS)	12 MS (VIS-NIR)
Ground sampling distance (nadir)	2.8 m PAN 5.6 m MS	100 m HS 30 m MS 150 m/300 m IR	6.5 m	22 m	5 m
Mission control center	United Kingdom	CAST China	Brandenburg, Germany	UK (and other)	Israel
Ground station	Fixed or mobile UK	CAST China	Svalbard, Norway	UK (and other)	Kiruna, Sweden

Notes: HS – hyper-spectral, PAN – multi-spectral, PAN – panchromatic, IR – infrared, VIS – visible, NIR – near infrared, UK – United Kingdom

1.6.1 Disaster Monitoring Constellation DMC-1

Small satellites provide a unique opportunity for affordable constellations. In this respect, small satellites can do things that are not practical with large satellites.

DMC may serve as the example for a constellation of five small satellites (DMC 2008). DMC has a GSD of 32 m and a swath width of 600 km (Landsat: GSD = 30 m, Swath width = 185 km). It provides a daily coverage of the Earth.

The five satellites

- AlSat-1,
- BILSAT-1,
- NigeriaSat-1,
- UK-DMC-1 and
- Beijing-1

from five countries have been launched with three COSMOS launchers into the same orbit. There are more constellations in sight, for instance DMC-2, a follow up of DMC-1 with improved performances based on new technologies, and RapidEye which was launched in August 2008 (see Fig. 1.5). The facts discussed in Sections 1.6.1–1.6.2 may lead to the conclusion: small satellites have the potential to change the economics of space and to increase the tempo of space exploitation.

DEIMOS-1

There are more satellites based on the DMC concept, for instance DEIMOS-1 of Spain. DEIMOS-1 is an automatic spacial platform with a small size and latest technology that provides optic and infra-red images adapted to the study of the earth’s vegetation cover.

			
<p>RapidEye Constellation of 5 spacecraft 7 year design lifetime 150kg</p> <p>High resolution multispectral imager, 6.5m GSD.</p> <p>To meet specific RapidEye GmbH. business needs</p>	<p>DMC+2.5 Constellation of 5 spacecraft 5-7 year design lifetime 160kg</p> <p>DMC and High resolution camera, 2.5m GSD. Hyperspectral and IR options</p>	<p>DMC+4 Designed to be part of constellation 5-7 year design life time DMC instrument and 4m GSD, 24km GSD panchromatic instrument.</p> <p>120-150kg depending on configuration</p> <p>Disaster monitoring and systematic mapping</p>	<p>TOPSAT derivatives Designed to be part of constellation 5-7 year mission design 2.5m GSD imager</p> <p>120-150kg depending on configuration</p> <p>Tactical Imaging and surveillance.</p>

Fig. 1.5 SSTL high resolution mission products (Curiel et al., 2005)

DEIMOS-1 is based on the concept Microsat-100 from Surrey. The satellite is conceived for obtaining Earth images with a good enough resolution to study the terrestrial vegetation cover, although with a great range of visual field in order to obtain those images with high temporal resolution and at a reduced cost. (Deimos-1, 2009)

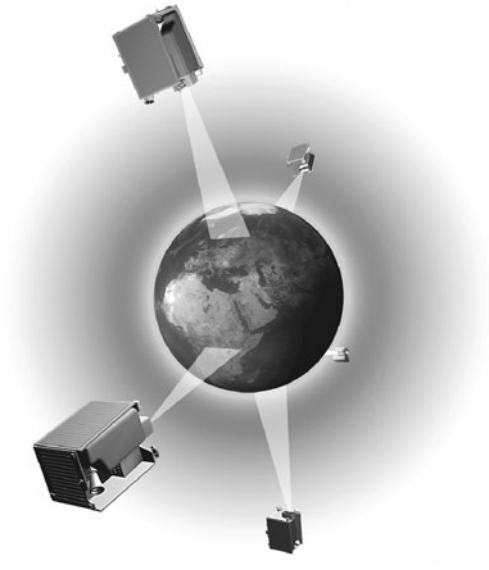
1.6.2 *RapidEye*

RapidEye is a commercial multispectral Earth observation mission of RapidEye AG of Brandenburg, Germany, that includes a constellation of five minisatellites. The mission will provide high-resolution multispectral imagery along with an operational GIS (Geographic Information System) service on a commercial basis. The objectives are to provide a range of Earth-observation products and services to a global user community. The RapidEye sensor images five optical bands in the 400–850 nm range and provides 6.5 m pixel size at nadir. It provides products for the following applications:

- Agricultural producers (farmers): Crop monitoring and mapping, yield prediction;
- Agricultural insurance: Provision of regularly updated field maps to help insurers assess insurance contracts and claims by providing quick and reliable information about damaged areas;
- Cartography – satellite based maps (scale 1:25,000), ortho photos, DEM (Digital Elevation Model) generation
- Other markets – disaster assessment, 3-D visualization
- Service spectrum at completion mission: Guaranteed daily revisit, global coverage, product delivery to the customer within 24 h, possibility of dedicated programming, capability of direct transmission and imagery transfer within hours, global digital database of “orthomaps” of 1:25,000 scale and DEMs of 20 m × 20 m resolution. The service permits also the merging of multi-temporal imagery with information from other sources.

The five RapidEye earth observation satellites have been launched on a single Russian Dnepr rocket from the Baikonur Cosmodrome in Kazakhstan in August of 2008. They are deployed in orbits at an altitude of 630 km. The satellites are placed equally spaced in a single sun-synchronous orbit to ensure consistent imaging conditions and a short revisit time. The satellites follow each other in their orbital plane at about 19 min intervals (Fig. 1.6). The constellation approach in a single orbital plane permits a cumulative swath to be built up (the spacecraft view adjacent regions of the ground, with image capture times separated by only a few minutes). A revisit time of 1 day can be obtained anywhere in the world ($\pm 70^\circ$ latitude) with body pointing techniques. The average coverage repeat period over mid-attitude regions (e.g., Europe and North America) is 5.5 days at nadir. The RapidEye system can access any area on Earth within 1 day and cover the entire agricultural areas of

Fig. 1.6 The RapidEye constellation consists of five satellites (DLR, 2006)



North America and Europe within 5 days. The swath width is 80 km and the maximum scene length per orb it is 1500 km. Some typical mission parameters of 5 small satellite missions are listed in Table 1.2.

1.7 Conclusion

Small Satellites for Earth observation supplement the conventional missions with advanced technologies and different instruments. Most of them, but not all, are cameras and imaging spectrometers in the visible and infrared wavelength range.

The fields of applications of the small satellite missions are very different but many of them are related to cartography, monitoring of environment changes, analysis and evaluation of natural disasters and environmental pollution, agriculture, remote sensing of land surface, land coverage and atmosphere including ionosphere.

Because of these application fields and the appropriate instrumentation a lot of small satellite missions are suitable for global change studies despite the fact that they are not originally designed for these purposes.

The first constellations of small satellites are operationally and assure a high area and time coverage. The instrumentation of these constellations is characterized by multi-spectral imaging systems. The data of these missions could be interesting for global change studies too. Small satellites make a constellation more affordable and for this reason in the future more constellations will be implemented using small satellites. They will be equipped not only with cameras and spectrometers but also with GNSS receiver and other instruments to study the atmosphere and ionosphere.

Today the data processing for small satellite missions runs on ground station in a very fast way to get the data product after 1 day or faster. In the future for certain applications small satellites will have the capability of on-board data processing and sending a high level data product (e.g. level 1b) to the final user without any delay.

References

- Background Paper (1998) Background Paper (No. 9). Small satellite missions – for the third United Nations conference on the exploration and peaceful uses of outer space. United Nations Document A/CONF. 184/BP/9, 26 May 1998
- Brieß K, Jahn H, Lorenz E, Oertel D, Skrbek W, Zhukov B (2003) Fire recognition potential of the bi-spectral infrared detection (BIRD) satellite. *Int J Remote Sens* 24(4):865–872
- DS Curiel et al (2005) http://www.dlr.de/iaa.symp/Portaldata/49/Resources/dokumente/archiv5/0301_daSilvaCuriel.pdf. Progress in small satellite technology for earth observation missions, da Silva Curiel, A., Cawthorne, A., Sweeting, Sir Martin, Accessed 25 June 2010
- Defense Update (2006) www.defense-update.com/directory/erosB.htm. EROS B Satellite Program. Accessed 25 June 2010
- DEIMOS-1 (2009) www.deimos-imaging.com/technology/our-satellite-deimos-1. Our satellite: DEIMOS-1. Accessed 25 June 2010
- DLR (2006) http://www.dlr.de/rd/desktopdefault.aspx/tabid-2440/3586_read-5336/. RapidEye. Accessed 25 June 2010
- EO (2001) http://directory.eoportal.org/get_announce.php?an_id=15557. TES (Technology Experiment Satellite). Accessed 25 June 2010
- ESA (2002) www.esa.int/esaMI/Proba_web_site/ESAIZ8NSRWC_0.html. Proba-1-Observing the Earth. Accessed 25 June 2010
- GEO (2008) www.earthobservations.org/. GEOSS on-line. Accessed 25 June 2010
- GMES (2009a) www.gmes.info. Gmes.info. Accessed 25 June 2010
- GMES (2009b) http://ec.europa.eu/gmes/index_en.htm. Gmes – Observing our planet for a safer world. Accessed 25 June 2010
- Kramer HJ (2002) Observation of the earth and its environment – survey of missions and sensors, 4th edn. Springer, Berlin
- Roeser HP, Eckardt A, v, Schoenermark M, Sandau R, Fricker P (2000) New potential and applications of ADS40. *Int Arch Photogramm Remote Sens XXXIII(B1)*: 251–257
- Sandau R (ed) (2006) International study on cost-effective earth observation missions. A. A. Balkema Publishers, a member of Taylor & Francis Group plc, Leiden, The Netherlands, ISBN 10: 0-415-39136-9, ISBN 13: 9-78-0-415-39136-8

Chapter 2

Spatial Pattern Analysis of Water-Driven Land Cover Change in Aridzone, Northwest of China

Qiming Zhou and Bo Sun

Abstract As a good indicator reflecting the interaction between human activities and ecological environment, landuse and land cover change has been studied for decades. The focuses of current researches are from “whether a change has happened” to “what cause the change and how will change in the future”. Many forces will drive environmental change of the earth. Water is one of the most significant factors for living things, especially in an arid environment. The study of land cover change with dynamic change of water resource can give us a better understanding on land cover change and the change reasons. This study attempts to find the spatial pattern of land cover change driven by water in an aridzone in the northwest of China and try to find a balance between economic development and sustainable development of environment for the fragile region. Since land cover change at that place is mainly due to the extension and abandon of farmland, two main land cover classes are considered in this study – the farmland and non-farmland. Based on the post-classification method, a change trajectory has been established to rebuild the history of land cover changes. In order to quantifying the impacts of water supply, two class-level metrics, Normalized Landscape Shape Index (NLSI) and Area Weighted Fractal Dimension Index (AWFDI), are adopted as tools which can measure the spatial distribution of changes. The result shows that farmland extension in that region is close to extreme due to the restriction of water supply. With the proposed approach, we can get a better understanding of the impacts of human activities or natural resources on land cover change.

Q. Zhou (✉)

Department of Geography, Hong Kong Baptist University, Kowloon Tong,
Kowloon, Hong Kong S.A.R, China
e-mail: qiming@hkbu.edu.hk

2.1 Introduction

As a land surface process, land use and land cover change (LUCC) emerged in researches on global environmental changes has been studied for several decades (Lambin et al. 2006). It was first recognized that land cover change modifies surface-atmosphere energy exchanges and thus influences global climate (Charney and Stone 1975; Otterman 1974). No matter how the theory of butterfly effect will affect this world, scientists believe that this change can make an impact on global change via carbon cycle, water cycle and so on (Turner et al. 1994; Woodwell et al. 1983). On the other hand, LUCC is not only a major part of global environmental change but also the cause of other changes at the same time. As the proportion of human behaviors in global change has been rising, land cover change can be regarded as a good indicator that reflects the interaction between human activities and natural environment (Li and Zhou 2009; Zhou et al. 2008c).

Since sensors on satellites can get a wide impression of landscape, remote sensing data, especially satellite images with different resolutions, are widely used for earth observation at different scales. One hot spot of global applications is land cover change detection. Coppin et al. (2004) divided existing methods of change detection into two broad classes, namely bi-temporal change detection and temporal trajectory analysis. To understand the trail of a rapid change, the latter is much more effective for modeling the dynamic process of changes (Zhou et al. 2008a, c) and forecast the trend of change in future (Tottrup and Rasmussen 2004; Liu and Zhou 2005). With the accumulation of remotely sensed images over the past decades, it is now possible to analyze the spatial pattern of land cover change over a long period.

Besides detecting changes, current research on LUCC is also focused on change driving forces and change estimation (Huang et al. 2009; Petit et al. 2001). In the analyses of change forces, a measuring tool called landscape metric is always employed (e.g., Crews-Meyer 2004, 2006). With the calculated metrics, spatio-temporal pattern of the changes can be better described and its relation to driving forces may be further explored (Zhou et al. 2008b).

With a fragile ecosystem aridzone is recognized as the sensitive region prone to rapid land cover change. This study, therefore, is focused on China's aridzone, where the environmental condition is dry and harsh, and rapid land cover change has been observed in the past decade due to the rapid growth of commercial agricultural operations. The irrigated farmland has been trebled, mainly by converting lands originally covered by natural rangeland vegetation into irrigated cotton fields. Serious questions, however, has been raised on whether this rapid growth can sustain given the intensified competition on water resources. When the supply of water become unsustainable, then the newly cultivated farmland will most likely be abandoned, leaving the sandy bare soil exposed to the strong seasonal wind without the protection of original rangeland vegetation.

This study aims to analyze land cover change process and to seek the relationship between land cover change and water resource. On the basis of trajectory analysis of land cover change, this study also attempts to get a quantified spatio-temporal description of the impacts of water resource on land cover change in the region.

2.2 Methodology

The methodology of this study is based on classification of multi-temporal images. It took the post-classification approach. Firstly, the multi-temporal images were classified to extract the farmland land cover type independently. Land cover change trajectories were then established to create categorical shift patterns between farmland and other land cover types. Finally, landscape metrics of land cover types were computed for analyzing the impact of water supply.

2.2.1 Study Area and Data

The study area is centered at Yuli County, Xinjiang Uygur Autonomous Region of China. It is located at the fringe of Taklimakan Desert. Topography shows a incline from higher terrain in the northwest to the southeast. This region represents a typical arid environment with desert vegetation and soils in the northwest of China. The rainfall in this region averages around 30 mm a year. Comparing with the total water consumption in Yuli County which reached to 0.25 billion m³ a year in recent years (Zhou 2007), the amount of rainfall is minimal. Tarim River, the largest inland river in China flows from the west to the east, while Konqi River flows out of Lake Bosten from the north, forming the major sources of water supply of the region. Dense vegetation is located along the rivers, forming a green corridor which is considered as one of the most important habitation area in the aridzone of China.

For a long history, the region had only limited human impact. With the economic reform in recent years, the national and local governments' policies have stimulated considerable investment on irrigated commercial cotton farming since early 1990s. According to the statistical yearbooks of Xinjiang (Statistics Bureau of Xinjiang Uygur Autonomous Region 1995, 2001, 2006, 2007, 2008), the farmland of the region grew from merely 9100 ha in 1994 to 35,500 ha in 2007, with the domination of cotton increased from 68% to over 99%.

This study used six multispectral remotely sensed images acquired in different years to establish a time series of land cover change, including multispectral images from Landsat-5/TM (25 September 1994), Landsat-7/ETM+ (17 September 2000), China-Brazil Environment and Resource Satellite (CBERS-02, 15 September 2005) and Beijing-1 micro-satellite (BJ-1, 10 August 2006, 31 August 2007 and 09 September 2008). The acquisition dates of these images are chosen to cover the rapid growing period of cotton crop, when a large contrast in multispectral data is shown to delineate the green vegetation from other land cover types. The 2005 CBERS image was registered and geo-referenced on 1:50,000 topographic maps. The other images were then geometrically corrected and registered on this 2005 master image. Efforts have been made to control the registration errors within half a pixel of the correspondent image so that the errors of change detection caused by mis-registration are less critical.

2.2.2 Image Classification

A supervised classification method using Maximum Likelihood classifier is employed to classify the images independently. Each image was classified into five or six classes independently to delineate land cover types. The five or six classes were then merged into only two classes, namely, farmland and the others. A post-classification process was applied to the unified classification results to remove isolated pixels using a majority filter with a 3×3 window. In order to make the image data comparable at the same spatial resolution so as to establish change trajectories at the pixel level, all the images need to be resampled to a 30 m resolution after classification.

The standard method of error matrices is used to assess classification accuracy for each image date. Both overall accuracy and Kappa index are computed as the accuracy assessment indices. The accuracies of these classified images are assessed independently. Due to the difficulty of getting historical reference data, accuracy assessments are conducted by visual inspection. For each classified image, more than 200 sample points including farmland and the others are randomly located on the original image with the reference to the field investigation data as the reference data set.

2.2.3 Establishment of Land Cover Change Trajectories

Trajectory is a typical snapshot model that records each situation of observed date during a time series. The focus of this study is the change of farmland in the study area. For our objective land cover types in this study, only two types are represented as the situations at each time point. Thus, the land cover trajectory in this research is simply defined as the situation of changes between farmland and the other land cover types. A trajectory can be specified as “farmland–others–farmland–others–others–others”, meaning that the land was once periodically cultivated, but finally abandoned. For a six-epoch, two-class scenario, the total number of possible trajectories is 64 (2^6). To establish the change trajectories, all classified images are integrated into a GIS with the raster format. After assigning value “1” (for farmland) or “0” (for others) to a corresponding bit position in a binary number for each pixel of each classified image, the multi-temporal images were merged together to identify every possible change trajectory with a unique number.

In this study, in order to reduce the risk of misunderstanding of trajectories which might be caused by roads and residents among farmlands, roads and residents are drawn with the help of high-resolution images (e.g., images from Google Earth) and then excluded when counting the farmland areas.

2.2.4 Spatial Pattern Analysis

As for a traditional irrigation system in that region, the amount of water supply impacts not only on the total area of farmland but also on the spatial pattern of farmland change. A quantified index that can describe the structure of landscape

is needed for modeling. Landscape metric can measure the geometric properties of landscape elements and their relative positions and distributions (Leitao et al. 2006). Two class-level metrics, namely Normalized Landscape Shape Index (NLSI) and Area Weighted Fractal Dimension Index (AWFDI), are computed as the indices of land cover change and used for further regression analysis.

These two parameters can be calculated in FRAGSTATS 3.3 by the following formulas (referred to Leitao et al. 2006; McGarigal et al. 2002):

$$\text{NLSI} = \frac{e_i - \min e_i}{\max e_i - \min e_i} \quad (2.1)$$

$$\text{AWFDI} = \sum_{j=1}^n \left[\left(\frac{2 \ln(0.25 p_{ij})}{\ln(a_{ij})} \right) \left(\frac{a_{ij}}{\sum_{j=1}^n a_{ij}} \right) \right] \quad (2.2)$$

where e_i = total length of edge of class i in terms of number of cell surface including all landscape boundary and background edge segments involving class i ; p_{ij} = perimeter (m) of patch ij ; a_{ij} = area (m^2) of patch ij .

To borrow the concepts of landscape when using the two indices, NLSI measures the aggregation of a class. NLSI equals 0 when the landscape consists of a single square or maximally compact, and 1 when the patch type is maximally disaggregated. AWFDI represents the shape complexity of a land cover type. The value of AWFDI ranges from 1 to 2. The closer to 1 the value is, the more complex the shape is.

Since the water supply for the study area in 2008 is much less than the same period of the preceding years (Xinjiang Tarim River Basin Management Bureau 2008), a comparison of land cover spatial pattern between 2008 and the preceding years is conducted as the proof that indicates how water supply impacts the spatial pattern of land cover change in that area.

2.3 Results and Analysis

2.3.1 Classification and Area Statistics

On the basis of the assessment on only two combined classes, the classifications have shown high accuracy between farmland and the others. According to the assessment result from those 200 samples, the overall accuracy of image classification ranges from 88.9 to 95.2% (from 1994 to 2008, they are 93.8, 95.2, 88.9, 90.7, 92.5 and 92.3% respectively), with Kappa coefficient ranging from 0.762 to 0.896 (from 1994 to 2008, they are 0.877, 0.896, 0.762, 0.799, 0.837 and 0.835 respectively). The result of classifications is satisfied in terms of the following change detection and spatial pattern analyses.

Table 2.1 illustrates the area statistics of farmland, which are compared with some reference data found in local statistical yearbooks issued by government. The reference data do not include some isolated administrative units in this study area,

Table 2.1 Comparison of area statistics between images and reference data

Year	Area statistics based on image classification		Reference data	
	Corps (ha)	% of study area	Cotton (ha)	% of total sown area
1994	12,200	3.2	9,100	68.2
2000	23,500	6.1	17,800	83.3
2005	34,400	9.0	28,000	97.1
2006	36,100	9.5	32,000	98.7
2007	43,500	11.7	35,500	99.3
2008	36,900	9.7		

Source of reference data: Statistics Bureau of Xinjiang Uygur Autonomous Region (1995, 2001, 2006, 2007, 2008).

so that some differences in the statistics between image and reference are expected. The results show that the sown area of farmland has rapidly increased till 2007 with average annual growth rate of 10.3%. Due to the limitation of water supply, the sown area of farmland in 2008 decreased 15.2% comparing with the former year of 2007.

2.3.2 Trajectories of Farmland Change

Figure 2.1 shows the result of land cover change trajectories. Since the land cover changes are mainly caused by increased farmland in the past decade with the

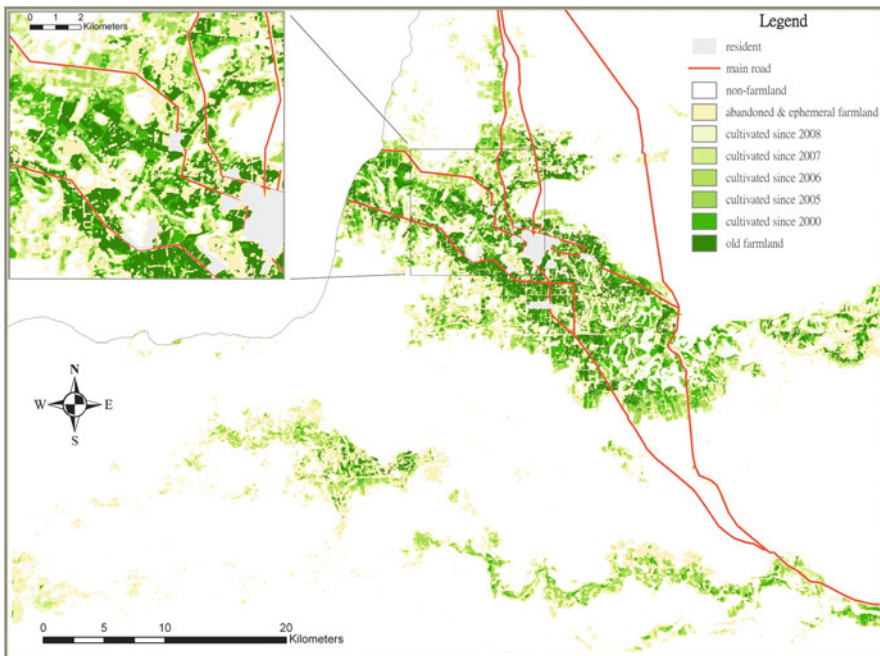


Fig. 2.1 Farmland change trajectories in Yuli County from 1994 to 2008

Table 2.2 Area statistics of farmland change trajectories

Trajectory	Description	Area (ha)	Total area (%)
X-X-X-X-X-X	Old farmland	6190	1.85
O-X-X-X-X-X	Cultivated since 2000	6130	1.83
O-O-X-X-X-X	Cultivated since 2005	6070	1.81
O-O-O-X-X-X	Cultivated since 2006	1840	0.55
O-O-O-O-X-X	Cultivated since 2007	2670	0.80
O-O-O-O-O-X	Cultivated since 2008	2610	0.78
X-O-O-O-O-O	Abandoned since 2000	280	0.08
X-X-O-O-O-O	Abandoned since 2005	130	0.04
X-X-X-O-O-O	Abandoned since 2006	50	0.02
X-X-X-X-O-O	Abandoned since 2007	140	0.04
X-X-X-X-X-O	Abandoned since 2008	510	0.15
X-O-X-O-X-X . . .	Ephemeral farmland	21,680	6.47
O-O-O-O-O-O	Non-farmland	28,680	85.59

Where cover type “X” = farmland; “O” = others.

exception of 2008, the increased farmland trajectories are highlighted, representing the old farmland since 1994 and the expansions since the other study periods. The trajectories are established without residents and main roads. Table 2.2 lists the calculated areas of farmland trajectories.

2.3.3 Impacts of Water Supply on Farmland Changes

Because we focus on the impacts of water resource on farmland change, the selected class-level metrics are adopted to represent the changes of farmland spatial distributions caused by water supply. The calculated metrics of farmland at each time points are listed in Table 2.3.

As shown in Table 2.3, the decreasing trend of NLSI during the whole study period illustrates an aggregation pattern in farmland extension that small patches are combined into big one. The change of AWFDI significantly reflects the change of water supply. The extension mode of farmland is a scattering mode in the fringe of older farmland when water supply is rich. However, once being lack of water like the situation in 2008, cultivated farmland is mainly concentrated along the river with a less complex edge, which shows a lower value of AWFDI.

Table 2.3 Metrics of farmland in different years

Year	NLSI	AWFDI
1994	0.1092	1.2037
2000	0.1179	1.2424
2005	0.1014	1.2517
2006	0.0819	1.2544
2007	0.0864	1.2844
2008	0.0667	1.2576

2.4 Discussion

With the constraint of image data acquisition, it is hard to acquire multitemporal images from the same platform in this study area. Post-classification comparison method has therefore been employed to cope with the multi-sensor multitemporal images. The principal advantage of the post-classification comparison method is that the images acquired on different dates by different sensors are independently classified, so that the problems of radiometric calibration among images of different dates by different sensors are minimized. However, It has been argued that the post-classification comparison method may overestimate land cover change due to uncertainties in data (Van Oort 2005), including those by errors in image registration and classification, and misunderstanding of trajectories.

Another issue in trajectory analysis is the abandoned farmland arose from new constructed residents and roads. The roads among farmland should be considered as the signature of a well developed farmland. It is not appropriate to put them into ephemeral trajectories. In order to minimize this kind of misunderstanding, roads and residents are excluded in the trajectory analysis part. From Table 2.2, it should be noted that the area of ephemeral farmland trajectories occupies 44.9% of total farmland trajectories. Compare with the pervious study, the percentage was just 23.5% (Zhou et al. 2008a, b). The reason is that a large number of cultivations in 2007 are abandoned in 2008.

Since the area of farmland has significant decreased in 2008 due to the decrease of water supply, it seems that farmland cultivation has reached the limit in this region. Unless much more significant measures are in place to increase efficiency of water use (e.g., the introduction of more advanced irrigation techniques), the situation of unbalanced demand and supply of water resource will be worse.

2.5 Conclusions

The quick response of remote sensing has been proved to be an effective tool for monitoring large-scale changes in earth observation. Land cover change happens widely and frequently. To understand the change itself better including change dynamics and driving forces through analyzing spatial patterns is in emergence. This paper has demonstrated a trajectory-analysis-based approach to model the spatio-temporal pattern of land cover change in an aridzone environment and analyzed the impacts of water supply on spatial and temporal pattern of land cover changes with measured metrics. As an improvement of those existing models which are conducted in solo aspect (e.g., trajectories of Normalized Difference of Vegetation Index in temporally; quantitative landscape metrics analysis without spatial relation), the proposed approach does not only consider rebuilding the process of land cover change but also describe the spatial distribution of the changes with quantified parameters.

According to this approach, the result shows that the farmland in that region has rapidly increased in the past decade. This reflects the impacts of local policy that

encourage cultivation for economy increasing. However, the conflict between farmland extension and the demand of water resource still exists. Farmland development has encountered the limit due to water supply. In order to better monitor land cover change and analyze the carrying capacity of cultivation in that region, an estimation model which can represent the relationship between water supply and farmland change is the goal of future work.

Acknowledgments The research is supported by National Key Basic Research and Development Program (2006CB701304), Research Grants Council Competitive Earmarked Research Grant (HKBU 2029/07P), and Hong Kong Baptist University Faculty Research Grant (FRG/06-07/II-76).

References

- Statistics Bureau of Xinjiang Uygur Autonomous Region (1995) Xinjiang statistical yearbook 1995 (in Chinese). China Statistics Press, Beijing
- Statistics Bureau of Xinjiang Uygur Autonomous Region (2001) Xinjiang statistical yearbook 2001. China Statistics Press, Beijing
- Statistics Bureau of Xinjiang Uygur Autonomous Region (2006) Xinjiang statistical yearbook 2006. China Statistics Press, Beijing
- Statistics Bureau of Xinjiang Uygur Autonomous Region (2007) Xinjiang statistical yearbook 2007. China Statistics Press, Beijing
- Statistics Bureau of Xinjiang Uygur Autonomous Region (2008) Xinjiang statistical yearbook 2008. China Statistics Press, Beijing
- Charney J, Stone PH (1975) Drought in the Sahara: a biogeophysical feedback mechanism. *Science* 187:434–435
- Coppin P, Jonckheere I, Nackaerts K, Muys B, Lambin E (2004) Digital change detection methods in ecosystem monitoring: a review. *Int J Remote Sens* 25(9):1565–1596
- Crews-Meyer KA (2004) Agricultural landscape change and stability in northeast Thailand: historical patch-level analysis. *Agric Ecosyst Environ* 101:155–169
- Crews-Meyer KA (2006) Temporal extensions of landscape ecology theory and practice: examples from the Peruvian Amazon. *Prof Geogr* 58(4):421–435
- Huang B, Zhang L, Wu B (2009) Spatio-temporal analysis of rural-urban land conversion. *Int J Geogr Inf Sci* 23(10) (Forthcoming)
- Lambin EF, Geist HJ, Rindfuss RR (2006) Introduction: local processes with global impacts. In: Lambin EF, Geist HJ (eds) *Land-use and land-cover change: local processes and global impacts*. Springer, Berlin
- Leitao AB, Miller J, Ahern J, McGarigal K (2006) *Measuring landscapes: a planner's handbook*. Island Press, Washington, DC
- Li B and Zhou Q (2009) Accuracy assessment on multi-temporal land cover change detection using a trajectory error matrix. *Int J Remote Sens* 30(5):1283–1296
- Liu H, Zhou Q (2005) Developing urban growth predictions from spatial indicators based on multitemporal images. *Comput Environ Urban Syst* 29:580–594
- McGarigal K, Cushman SA, Neel MC, Ene E (2002) FRAGSTATS: spatial pattern analysis program for categorical maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at the following web site: www.umass.edu/landeco/research/fragstats/fragstats.html. Accessed 12 September 2009
- Otterman J (1974) Baring high-albedo soils by overgrazing: a hypothesized desertification mechanism. *Science* 86:531–533
- Petit C, Scudder T, Lambin, E (2001) Quantifying processes of land-cover change by remote sensing: resettlement and rapid land-cover changes in south-eastern Zambia. *Int J Remote Sens* 22(17):3435–3456

- Tottrup C, Rasmussen MS (2004) Mapping long-term changes in savannah crop productivity in Senegal through trend analysis of time series of remote sensing data. *Agric Ecosyst Environ* 103:5445–560
- Turner MG, Meyer WB, Skole DL (1994) Global land-use/land-cover change: towards an integrated study. *Ambio* 23:91–95
- Van Oort PAJ (2005) Improving land cover change estimates by accounting for classification errors. *Int J Remote Sens* 26(14):3009–3024
- Woodwell GM, Hobbie JE, Houghton RA, Melillo JM, Moore B, Peterson BJ, Shaver GR (1983) Global deforestation: contribution to atmospheric carbon dioxide. *Science* 222:1081–1086
- Xinjiang Tarim River Basin Management Bureau (2008) Report on water discharge information of Tarim River Basin from October 2007 to Mid-July 2008 (in Chinese). Xinjiang Tarim River Basin Management Bureau website. <http://www.tahe.gov.cn/e/action/ShowInfo.php?classid=65&id=8554>. Accessed 31 May 2010
- Zhou H (2007) Study on landscape ecology of river corridor in aridzone: to take the middle and lower reach of Tarim river region, Xinjiang for example (in Chinese). Science Press, Beijing
- Zhou Q, Li B, Kurban A (2008a) Trajectory analysis of land cover change in arid environment of China. *Int J Remote Sens* 29(4):1093–1107
- Zhou Q, Li B, Kurban A (2008b) Spatial pattern analysis of land cover change trajectories in Tarim Basin, northwest China, *Int J Remote Sens* 29(19):5495–5509
- Zhou Q, Li B, Sun B (2008c) Modelling spatio-temporal pattern of landuse change using multi-temporal remotely sensed imagery. In: Chen J, Jiang J, van Genderen J (eds) *The International Archives of The Photogrammetry, Remote Sensing and Spatial Information Sciences* 37(B7):729–734

Chapter 3

A Spatial Logistic Regression Model for Simulating Land Use Patterns: A Case Study of the Shiraz Metropolitan Area of Iran

Amin Tayyebi, Mahmoud Reza Delavar, Mohammad Javad Yazdanpanah, Bryan Christopher Pijanowski, Sara Saeedi, and Amir Hossein Tayyebi

Abstract Changes in global land cover and land use are occurring at a rate, magnitude, and spatial extent unprecedented in human history. Land use and land cover change is one of the most important components and major causes to global environmental changes. This paper presents an urban expansion model which utilizes Logistic Regression (LR) as a means to simulate and predict urban expansion pattern. LR is used as the core algorithm in this model and remote sensed image with arbitrary time interval and environmental variables have been employed in geospatial information system. Socio-economic and environmental variables were used as inputs while urban and non-urban areas were considered as outputs for LR model. Evaluation of the model was performed with Relative Operating Characteristic (ROC). A case study of urban expansion in Shiraz Metropolitan Area (SMA) is presented to simulate land use change. This paper presents a version of LR model which is parameterized for SMA and explores how factors such as road, building area, service centre, green space, elevation and slope can influence urban expansion. The area under the operating characteristic curve produced an accuracy of 68%. Having model parameters with specific time interval and assuming the existence of the same rate of urban expansion, we used the LR model presented here to generate locations of future urban expansion.

3.1 Introduction

Humans impact the environment in a variety of ways. Stresses to natural resources occur through both the continuing increase in size of the human population and exacerbated by a rapid development of technologies that exploit the plant's natural resources. Global change research analyses the natural variability of the Earth

A. Tayyebi (✉)
Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran
e-mail: amin.tayyebi@gmail.com

System and the causes, mechanisms and effects of the complex interactions between its components and the human population. It develops procedures that make the distinction between natural variability and human-induced changes as well as create tools that allow communities prognostic abilities to foresee changes and their consequences, to identify options for counteraction and to balance adaptation to change and mitigation of causes. Both global and regional strategies have to be developed in order to balance the interaction between humans and the Earth System and at the same time to ensure a sustainable development of human societies. It is at the local and regional levels where decisions that promote sustainable development are most likely to be implemented and where action occurs. The focus on regions and common efforts to bridge the gap between global perspectives on change as well as the analysis of regional impacts and the development of regional sustainable management options must therefore be intensified.

Land use/land cover change (LUCC) is driven by the interaction in space and time between biophysical and human factors. Urbanization is a rapid land use change process that produces different patterns depending on the proximity to large urban cities across the landscape (Wu, 2004). LUCC results from the complex interaction of social, ecological and geophysical processes. An urban land use system is dominated by human activities with complex spatio-temporal dynamics. Traditionally, the global change research community has produced three kinds of spatial land use change models: empirical estimation models, dynamic simulation models and rule-based simulation models (He and Lo, 2007). Dynamic simulation models and empirical estimation models have been used to model land use changes. Rule-based simulation models are most suitable for incorporating spatial interaction effects and handling temporal dynamics. However, they focus on simulation of spatial pattern rather than on interpretation or understanding of spatio-temporal processes of urban growth. Most dynamic simulation models cannot incorporate enough socioeconomic variables. Logistic regression (LR) models quantify the relationship between drivers and probability of land use change. Many land use change models have been used to project future land use changes in urban areas based on past trends and the drivers thought to determine conversions of land between different categories.

Information about urbanization, as analyzed from multiple time periods, can generate useful knowledge about the patterns of urban and the possible factors driving the changes. This information has been used by planners and resource managers to make better decisions that affect the environment and local and regional economies. Decision makers are increasingly relying on models of LUCC; therefore, users need appropriate ways to interpret the models and to communicate the information that LUCC models produce (Veldkamp and Lambin, 2001; Veldkamp and Verburg, 2004; Verburg and Veldkamp, 2004). Description and modelling of land systems highly depend on the data availability and quality. Remote Sensing (RS) and Geospatial Information System (GIS) provide us with an efficient spatial capability to monitor urban expansion in urban areas. Land Transformation Model (LTM) is a land use change model that uses Artificial Neural Networks (ANNs) and GIS (Pijanowski et al., 2000, 2002, and 2005). Monitoring urban expansion and land use change based on multi-temporal GIS maps and remote sensing images

were performed in Iran, respectively (Tayyebi, et al., 2008a, 2008b; Pijanowski et al., 2009). Empirical estimation models use statistical techniques to model the relationships between land use changes and the drivers based on historic data. As an empirical estimation method, LR has been used in deforestation analysis (Geoghegan et al., 2001; Schneider and Pontius, 2001), agriculture (Serneels and Lambin, 2001), and urban growth modeling (Allen and Lu, 2003; Landis and Zhang, 1998; Wu and Yeh, 1997). In many cases, these models fit spatial processes and land use change outcome reasonably well (Irwin and Geoghegan, 2001). Allen and Lu (2003) also highlighted use of statistical methods, mainly LR, to identify which variables best represent urban change process properly. Cross tabulation matrices are used to quantify the association between simulated and reference map (Pontius et al., 2004a, 2004b; Pontius and Spencer 2005; Tayyebi et al., 2009a, 2009b). This has been extended recently to include eight different kinds of calibration metrics used to estimate model goodness-of-fit: four location-based measures and four patch metrics (Pijanowski et al. 2006).

This paper presents an overview and application of an urban expansion model which take advantage of GIS, LR and RS to simulate and predict urbanization in Shiraz Metropolitan Area (SMA) of Iran. Predictor variables which are usually important landscape features were used as inputs while urban and non-urban areas were used as outputs in LR model. Two Landsat TM images of SMA in 1988 and 2000 were rectified and registered to Universal Transverse Mercator (UTM) WGS 1984 zone 39 N. Supervised classification was used to classify the images to different land use categories. In addition three and two land use classes that were extracted from the first and the second satellite images respectively, service centres were obtained from county road maps stored as point coverage; two other classes were added from a topographic data at a scale of 1:50,000. Our database was included these eight classes: road, build-up area, service centre, green space, elevation, slope, urban and non-urban areas. Having model parameters between 1988 and 2000 and assuming the existence of the same rate in urban expansion, new urban areas of SMA has been derived for 2012. The proposed model was evaluated with respect to spatial and temporal sample datasets. The accuracy of the model was 68% with ROC metrics consideration to simulate pattern of urban areas.

The structure of the rest of the paper is as follows. In following section, methodology introduces different steps of LR model and the underlying assumptions in urban land use change. The next section provides a brief summary of study area, data sources and presents how the LR is parameterized for SMA. The Results section provides a summary of goodness of fit metrics for SMA and our conclusion highlights future challenges and directions for LR.

3.2 Methodology

LR model was used to associate the urban growth with demographic, econometric and biophysical driving forces and to generate an urban expansion probability map. In a raster GIS modeling environment, the data layers are tessellated to form a grid of cells. The nature of the LUCC of a cell is dichotomous: either the presence of urban

growth or absence of urban growth. If binary values 1 and 0 are used to represent urban growth and no urban growth respectively. Main stages of LR model are as follows: (1) Data preparation; (2) Simulation process; (3) Model Evaluation and (4) Prediction.

3.2.1 Data Preparation

The first and the second satellite images were geometrically registered to the UTM projection system. Polynomial transformation function was used to resample the projected images to the same resolution. Then both satellite images were classified to different land use categories. These classes were converted to separate spatial layers that can be used as input and output of the model. LR as a mathematical equation was flexibly designed to have different spatial or non-spatial layers. On the second satellite image, urban and non-urban areas were identified as output of LR. Then, other classes of the required topographic data were added to the database. Coding of data to create spatial layers of predictor variables are generated from a series of base layers that are stored and managed within a GIS. Spatial or non-spatial function relates input variables to LR model. Three classes of functions including neighbourhoods (densities) function; site specific characteristics function and distance function from the location of a predictor cell have been used (Pijanowski et al., 2002).

3.2.2 Simulation Process

In the application of LR, each observation is a grid cell. The dependent variable is a binary presence or absence event, where 1 = urban areas and 0 = non-urban areas. The function is a monotonic curvilinear response bounded between 0 and 1, given by a logistic function of the form (Eq. (1)) (Pontius and Schneider, 2001):

$$P = E(Y) = \frac{\exp\left(\beta_0 + \sum_{i=1}^n \beta_i X_i\right)}{1 + \exp\left(\beta_0 + \sum_{i=1}^n \beta_i X_i\right)} \quad (1)$$

The logistic function gives the probability of urban expansion as a function of the explanatory variables.

Where:

P: probability of urban expansion in the cell

E(Y): expected value of the binary dependent variable Y

X_i : predictor variables (socio-economic and environmental variables)

β_0 : constant to be estimated

β_i : coefficient to be estimated for each independent variable X_i

The logistic function can be transformed into a linear response with the transformation (Pontius and Schneider, 2001):

$$P' = \log_e \left(\frac{P}{1-P} \right) \quad (2)$$

Hence

$$P' = \beta_0 X_i + \sum_{i=1}^n \beta_i X_i \quad (3)$$

The transformation (Eq. (2)) from the curvilinear response (Eq. (1)) to a linear function (Eq. (3)) is called a logic or logistic transformation (Pontius and Schneider, 2001). The transformed function allows linear regression to estimate each β_i . Since each of the observations is a cell, the final result is a probability score (P) for each cell. The output of LR is a map of likelihood values, which specifies the relative likelihood of change for each cell based on the predictor variable values.

3.2.3 Model Evaluation

Relative Operating Characteristic (ROC) was used to validate the LR model. Recently the ROC method was brought to the field of LUCC modeling to measure the relationship between simulated change and real change (Pontius, 2000; Schneider and Pontius, 2001). ROC method is an excellent method to evaluate the validity of a model that predicts the occurrence of an event by comparing a probability image depicting the probability of that event occurring and a binary image showing where that class actually exists (Pontius 2002; Tayyebi et al., 2009a, 2009b; Pijanowski et al., 2009) created when simulated and observed maps are compared. Model validation using ROC reported a summary ROC value, a ROC curve as well as the coordinates of the points on the curves that were used to calculate the ROC value. ROC curve plots the rate of true positive to positive classifications against the rate of false positive to negative classifications as threshold value is varied between 0.0 and 1.0. A non-parametric approximation using SPSS (SPSS Inc, 2003) is used here to estimate the area under the curve produced by varying the threshold and plotting numerous thresholds and resultant values of specificity and sensitivity.

3.2.4 Prediction

The probability map can be used for producing maps of urban distribution if any quantitative data on the future total areas of urban distribution, for example, urban planning data, are given. In prediction process, after the parameters of the mathematical model were obtained and the LR model was tested successfully, the

feed-forward algorithm was used in the LR model with new topology. In this topology, spatial layers extracted from the second satellite image (predictor variable) were replaced with spatial layers from the first satellite image as input for the LR in the prediction process. The testing process that followed used driving variable grids from all cells in the study locations but with the output values removed. Output of LR in this phase means the future urban expansion for that region. The critical threshold value is the lowest change likelihood value selected for cells to transition during the period.

3.3 Parameterization of SLR for Shiraz Metropolitan Area

This section presents a complete description for LR model starting from image pre-processing, registration, image classification and constructing the LR model. Input layers represent phenomena which may influence the LR model. It is found that six independent variables influence urban expansion in SMA: slope, elevation, distance from building area, distance from service centre, distance from green space and distance from road. Each independent variable map is standardized between 0 and 1 that indicates the relative suitability at a value of the independent variable.

3.3.1 Study Area and Data Sources

Shiraz, a metropolitan area located the center of the Fars province of Iran, is located at Longitude 32–52'E, Latitude 29–37'N and 1540 m above the sea level in the south of Iran (Fig. 3.1). It is 895 km away from Tehran and its area is 220 km². Shiraz is the place of new jobs, recreational facilities and beautiful buildings that attract many migrants.

Over the last 50 years Shiraz has experienced tremendous in-migration of people from rural areas and small cities. According to the 2005 census, Shiraz's population was estimated at 1,442,842, having increased from 848,289 in 1985. Table 3.1 summarizes trends of Shiraz population from 1985 to 2005.

3.3.2 Implementing the SLR

The Landsat images were geometrically registered to the Universal Transverse Mercator (UTM) WGS 1984 zone 39 N. Registration errors were about 0.52 pixels. SMA was extracted from the two Landsat images in order to have less process and time saving for classification. The first image (1988) and the second one (2000) were subjected to a classification of zones. Supervised classification was utilized to classify the images to different land use categories. All land use classes of SMA were also reclassified from their original classification to Anderson Level I (Anderson et al. 1976) for the simulating exercises. In order to classify the registered images,



Fig. 3.1 Map of SMA

Table 3.1 Trends of Shiraz population from 1985 to 2005

Year	Population number
1985	848,289
1995	1,053,025
2005	1,442,842

three classes of interest were selected from different classes in the images namely: road, building area and green space. In addition, locations of service centres were obtained from county road maps at a scale of 1:50,000 and stored as point coverage. The accuracy of testing process for the classification of Landsat TM image of 1988 was 83.42%, while it was 86.12% for Landsat TM image of 2000. Kappa standard is 0.771 for Landsat TM image of 1988 while 0.792 for that of 2000. These parameters were input to ArcMap software for calculations which were different due to type of parameters. After these calculations different layers have been stored in grid format then each location contained its spatial configuration value from each driving variable grid. Briefly, each value in an entire driving variable grid was normalized from 0.0 to 1.0 by dividing each value by the maximum value contained in driving variable grid. Effective parameters require main considerations and criteria listed as follows.

3.3.2.1 Absorbing Excursion Spaces

Denser urban areas which already have access to utilities like water lines, sewer lines, electricity, and cable services among others also can be expected to attract residential development because of the reduced cost of connecting to these services (Pijanowski et al., 2005). Areas closest to urban areas are likely locations for urbanization because they are assumed to require the least monetary cost for connecting to urban services such as water and sewer (Pijanowski et al., 2005). Here, we consider absorbing excursion spaces as distance from service centre, green space and build-up area. The distance of each cell from the nearest absorbing cell was calculated and stored as a separate variable grids. These variable grids represented the potential effect of a location for urban expansion.

3.3.2.2 Transportation

It is widely accepted in LUCC literature that the building of roads will spur development because of improved accessibility. Transportation, another important factor, is the distance of each cell from the nearest road cell calculated and stored in separate Arc/Info Grid coverage. The hypothesis is that humans need roads to access areas where resources will be used resulting in urban change.

3.3.2.3 Landscape Feature

Landscape topography is an influential factor contributing toward build-up area utilization. Elevation is important in flood prone areas. Slope is important to developers who want to minimize landscape costs. Then, for each cell in the study area, there is a vector of 6 by 1 measurements as input of LR model. Figure 3.2 shows six variables compiled in Arc/Info Grid format as inputs of LR model at 1988.

There are two constraints for simulation of new urban areas in SMA. First, cells that are urban areas in 1988 are obviously not candidates for new urban areas in 2000. Second, cells that are protected legally from new urban areas are assigned the absolute lowest suitability value in the final suitability maps.

The second Landsat image was subjected to a classification of two zones (urban and non-urban areas). The output vector was coded to represent cells as binary variable which values from 0 (non-urban) to 1 (urban). Figure 3.3 shows this variable as output of LR model compiled in Arc/Info Grid format at 2000.

3.4 Results and Discussions

Study area includes 245,588 cells which 85,956 (35.0%) of the cells have limitation to undergo transition while 159,632 (65.0%) of the cells can be subjected to transition in SMA. LR model estimated 127,706 (80.0%) of the qualified cells had change their likelihood values to 0 while 3,193 (2.0%) had likelihood their values to 1 and other cells 28,733 (18.0%) have value between 0.0 and 1.0. Cells with values closest

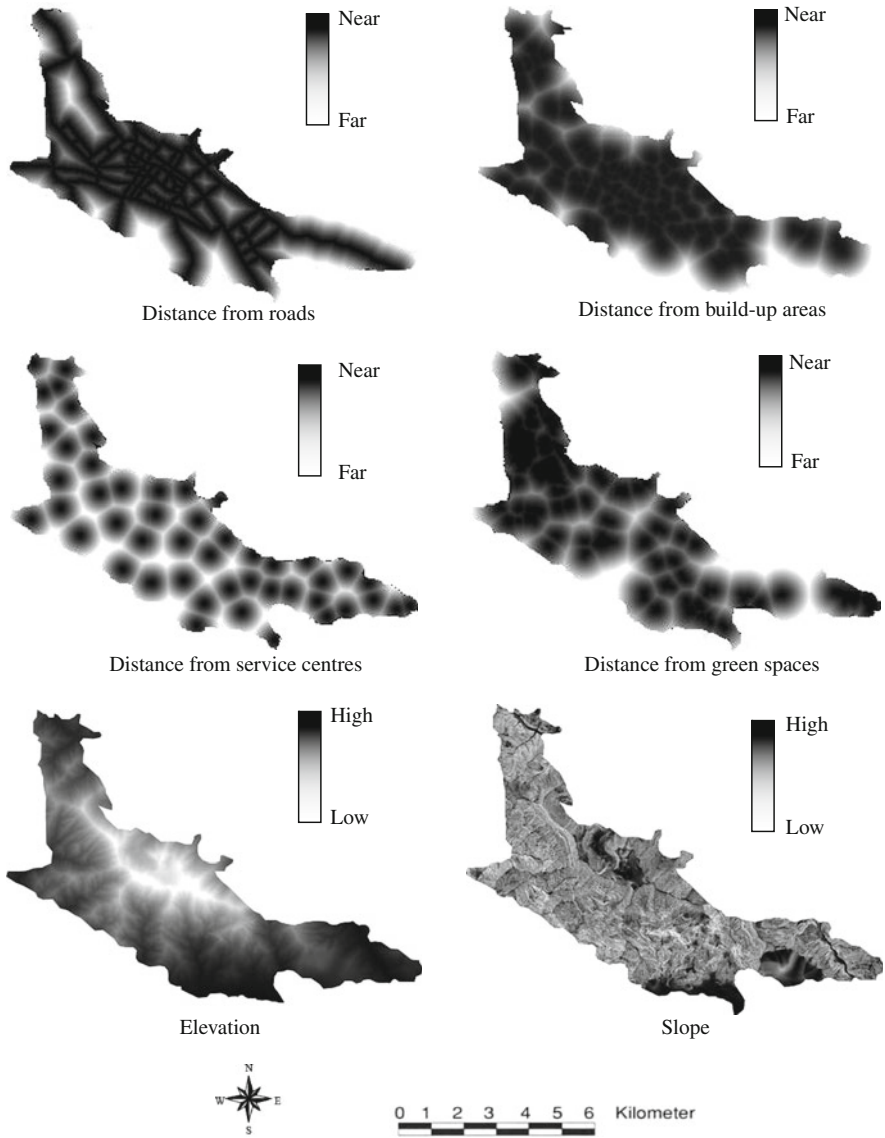
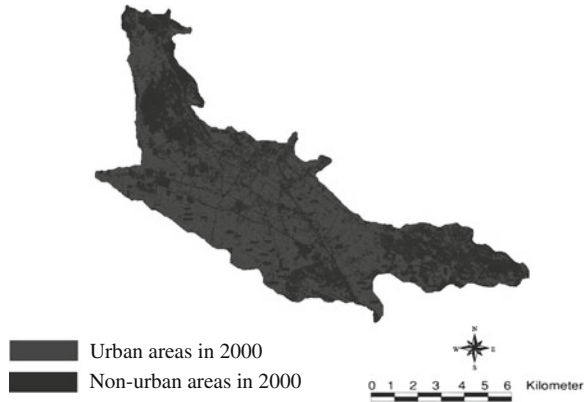


Fig. 3.2 SMA maps of the six variables used as input for the LR in 1988

to 1 were selected as locations most likely to transition. The critical threshold value, which is the lowest change likelihood value selected for cells to transition during the 12-year period, was 0.5. Only 3.5% of all qualified areas changed to urban area in the observed databases. Results show that only 5,587 cells undergo the transitions in SMA.

Fig. 3.3 SMA maps of the two-variables used as output for the LR in 2000



3.4.1 Model Coefficients

Pontius and Schneider (2001) explain how to use the ROC technique to examine how well a suitability map portrays the true locations in a Boolean image—for our case the Boolean image represents actual urban expansion between 1988 and 2000. LR as a mathematical equation was used for simulating and prediction process. LR model was designed to have a flexible number of inputs depending on the number of variables presented to it (Table 3.2).

A stratified random sampling method was used to extract sample subset from the complete coverage of 1988–2000 data to train LR model. The stratified sampling method assures that the generated sample subsets represent different urban expansion patterns in the region. All inputs were normalized to a range from 0.0 to 1.0. The patterns file contained information from the six input grids and output file for LR model so that each line in the pattern file corresponded to one location. The output layer for LR model contained binary data represented whether a cell location changed (1 = urban; 0 = non-urban) during the study period (Pontius and Schneider, 2001) (Eq. (4)). To avoid over-training of LR model, the parameters of the mathematical equation was obtained with a partial set of data by providing it with data from every other cell in the city.

Table 3.2 Parameters of LR model in SMA

Variable	Description	Nature of variable
X_1	Distance from road	Continuous
X_2	Distance from build-up area	Continuous
X_3	Distance from service centre	Continuous
X_4	Distance from green space	Continuous
X_5	Elevation	Continuous
X_6	Slope	Continuous
P	Probability of urban expansion	Discrete

Table 3.3 The results of LR parameters

Variable	Urbanland	
	β	Exp(β)
Constant	0.265	1.303
X_1	0.001	1.001
X_2	0.066	1.068
X_3	-0.071	0.931
X_4	0.002	1.002
X_5	0.062	1.064
X_6	-0.059	0.942
ROC		0.680

$$P = E(Y) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6)} \quad (4)$$

The results of LR for different land use are summarized in Table 3.3.

3.4.2 Model Evaluation

Cells that were simulated as transition to new urban areas were compared with the cells that actually did transition during the period of study. The ROC compares binary data over the whole range of simulated probabilities. It aggregates into a single index of agreement, reflecting the ability of the model to predict the probability of urban land distribution at various locations on the landscape. The ROC is a measure of the ability of the model to correctly specify location across a range of threshold values. For our case the Boolean image represents actual development between 1988 and 2000 in testing process. The file generated from the training process was used to estimate output values for each location in testing process. The testing process that followed used driving variable grids from all cells in the study locations between 1988 and 2000 but with the output values removed. When the map of a simulated urban expansion in 2000 is overlaid on the reference map 2000, a contingency table is used for accuracy assessment in testing process. Cells that were simulated as transition to new development areas were compared with the cells that actually did transition during the period of study.

First the ranked image of probability of urbanization was sliced at a series of threshold levels (He and Lo, 2007). A threshold refers to the percentage of cells in the probability image to be reclassified as 1 in preparation for comparison with the reference image. The series of thresholds was specified at an equal interval of 5%. The threshold values are cumulative, therefore setting the equal interval thresholds 5, 10, 15...95 would yield 20 threshold intervals 0–5%, 0–10%, 0–15%, ... and 95–100%. ROC began with the cell ranked the highest for probability, reclassified it as 1 and continued down through the ranked cells until 5% of the cells had been reclassified as 1. The remaining 95% was classified as 0. This slice image was then

Table 3.4 Contingency table showing the comparison of the slice image of predicted urban expansion probability with the reference image

		Reference image	
		Urban expansion (1)	No urban expansion (0)
<i>Slice image of predicted probability</i>	Urban expansion (1)	A (True Positive)	B (False Positive)
	No urban expansion (0)	C (False Negative)	D (True Negative)

compared with the reference image. Then ROC continued for the successive threshold. For each slice generated from each threshold, a two-by-two contingency table was created based on the comparison of the slice image with the reference image (Table 3.4). In the table, A represents the number of true positive cells which are predicted as urban expansion and are actually urban expansion in the reference image. B is the number of false positive cells. C is the number of false negative cells. D is the number of true negative cells. From each contingency table for each threshold, one data point (x, y) was generated where x (Eq. (5)) is the rate of false positives (false positive percentage) and y (Eq. (6)) is the rate of true positives (true positive percentage):

$$\text{True Positive\%} = \frac{A}{A + C} \quad (5)$$

$$\text{False Positive\%} = \frac{B}{B + D} \quad (6)$$

These data points were connected to create a ROC curve from which the ROC value was calculated. The ROC statistic is the area under the curve that connects the plotted points. The ROC curve is shown in Fig. 3.4. The ROC value is 0.68.

3.4.3 Forecasting

To produce the spatial pattern of urban distribution given a certain amount of urban area, the increase of the number of urban cells compared to the 1988 base urban map was calculated. Then the number of urbanized cells was allocated to the probability map in the order of high probability value to low probability value. This generated a growth map. Next the growth map was combined with the 1988 base map to produce the urban distribution map. After LR model was trained and tested successfully, the parameters were obtained and the feed-forward algorithm was used for prediction. Assuming the same rate of urban expansion in SMA during 1988 and 2000, the predicted urban expansion for the year 2012 based on the full dataset of urban expansion in the year 2000 was determined. Figure 3.5 shows urban expansion in SMA in 2012 which is overlaid on SMA in 1988 and 2000.

Future urban expansion appears to be focused at the east of SMA. A great deal of clumped development is anticipated in the North West and South East portion of SMA. In addition, a great deal of dispersed development is anticipated in the

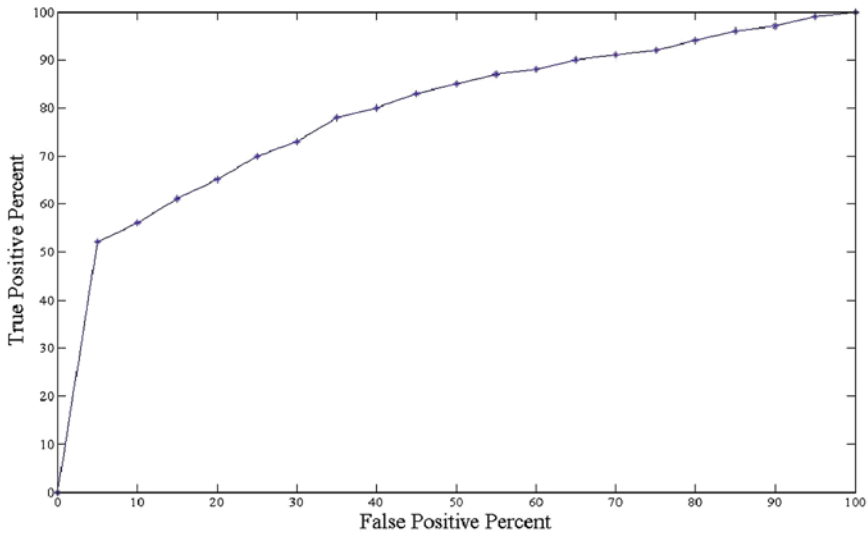


Fig. 3.4 ROC curve

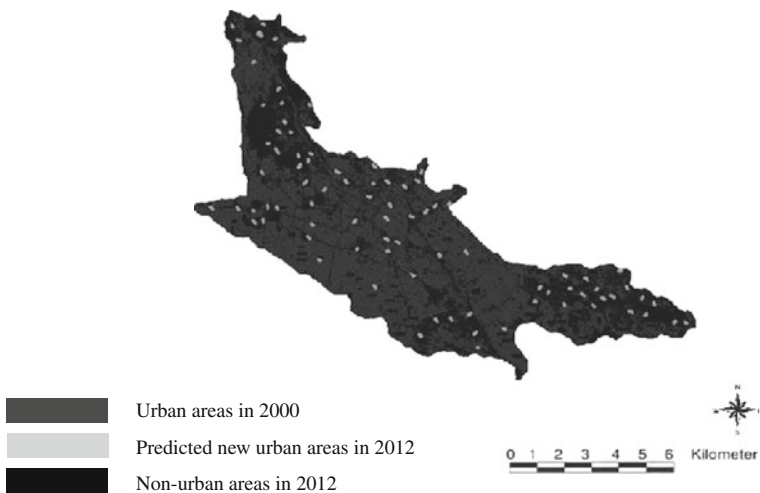


Fig. 3.5 Prediction of LR in 2012 for SMA

South portion of the SMA. But, there is nearly no development into the centre and South West of SMA predicted. Because in centre of SMA, there is no space that can be developed and in South West of SMA, there is a legal restriction for development from the government side. What is interesting from these forecasting results is that the model predicts new urban areas to be occurred along build-up areas. The

Table 3.5 ROC values in simulation process with reduced-variable

Reduced-variable	ROC value
Distance to road	0.48
Distance to build up area	0.53
Distance to green space	0.55
Slope	0.59
Distance to service centre	0.63
Elevation	0.67

model predicts that these urban expansion patterns will be different for various parts of SMA.

For the rank order of input variables according their influence on the model performance, we compared the predictive ability of the six versions of the reduced-variable model of input variables. We excluded each variable and perform LR model with other six-variables respectively. We saved the ROC for each of the six-variable models. The resultant value represents the relative effect, of each predictor variable on the model performance. The ROC values of the six-variable models were compared against each other and it was concluded that the rank order of predictor variables according their influence on the model performance was respectively: distance from road, distance to building areas, distance from green space, slope, distance from service centre and elevation (Table 3.5). When distance to road parameter is removed in simulation process, ROC value has maximum decline. When elevation parameter is removed in simulation process, ROC value has minimum decrease, so, elevation parameter has no effect in simulation process.

3.5 Conclusions

Urban expansion model has been implemented in this paper which takes advantage of GIS, LR and RS based on the utilization of a variety of social and environmental factors. LR modeling was used to identify and improve our understanding of the demographic, econometric and biophysical forces that have drive urban expansion and to find the most probable sites of urban expansion in SMA of Iran. LR model examines the relationship between predictor variables as inputs and urban/non-urban areas as outputs to model urban expansion. This linkage is based on the fact that land use and land cover data are the main input parameters for urban expansion analyses. Integration LR with GIS is essential for modelling urban changes because of the spatial nature of many the input variables. This model also allows the user to block out areas where development cannot occur because of either strict policy enforcement or physical constraints.

Applying the LR model to SMA allowed us to examine impacts of different variables on model performance. The following groups of factors were found to affect urban expansion in different degrees: distance from road, distance from build-up area, distance from service centre and distance from green space, elevation and

slope. With developing the six versions of LR model, each with one of the variables removed, the relative contributions of each variable on model performance were assessed. Distance to roads was the best predictor of urban expansion during the 1988–2000 periods in SMA. ROC metric have been employed to determine how well LR performed on the study area. A future simulated map driven by the six variables resulting from LR forecasts is compared with real map provide a satisfactory means for predicting change inside of dense urban areas and urban fringes for urban planning. A map of urban expansion probability was calculated and used to predict future urban patterns. ROC value of 68% indicates that the probability map is valid. The LR model predicts that these growth patterns will be different for various parts of SMA. Not surprisingly, it that forecasted results shows that the model predicts that urban expansion occurs along build-up areas. It was concluded that despite LR's lack o temporal dynamics, it was spatially explicit and allowed much deeper understanding of the forces driving the growth and the formation of the urban spatial pattern.

References

- Allen J, Lu K (2003) Modeling and prediction of future urban growth in the Charleston region of South Carolina: a GIS-based integrated approach. *Conserv Ecol* 8(2):2 [online]. URL:<http://www.consecol.org/vol8/iss2/art2>
- Anderson JR, Hardy EE, Roach JT, Witmer RE (1976) A land use and land cover classification system for use with remote sensor data. US Geological Survey, Professional Paper 964: 28, Reston, VA
- Geoghegan J, Villar SC, Klepeis P, Mendoza PM, Himmelberger YO, Chowdhury RR (2001) Modeling tropical deforestation in the southern Yucatan peninsula region: comparing survey and satellite data. *Agric Ecosyst Environ* 85:25–46
- He Z, Lo C (2007) Modelling urban growth in Atlanta using logistic regression. *Comput Environ Urban Syst* 31(6):667–688
- Irwin EG, Geoghegan J (2001) Theory, data, methods: developing spatially explicit economic models of land use change. *Agric Ecosyst Environ* 85:7–23
- Landis J, Zhang M (1998) The second generation of the California urban futures model. *Environ Plann B: Plann Des* 25:795–824
- Pijanowski BC, Gage SH, Long DT (2000) A land transformation model: integrating Policy, Socioeconomic and environmental drivers using a geographic information system. In: Sanderson J, Harris L (eds) *Landscape ecology: a top down approach*. CRC Press (Lewis), Boca-Raton, FL
- Pijanowski BC, Brown DG, Shellito BA, Manik GA (2002) Using neural networks and GIS to forecast land use changes: a land transformation model. *Comput Environ Urban Syst* 26(6):553–575
- Pijanowski BC, Pithadia S, Shellito BA, Alexandridis K (2005) Calibrating a neural network-based urban change model for two metropolitan areas of Upper Midwest of the United States. *Int J Geograp Info Sci* 19:197–215
- Pijanowski BC, Alexandridis KT, Muller D (2006) Modeling urbanization patterns in two diverse regions of the world. *J Land Use Sci* 1(2–4):83–109
- Pijanowski BC, Tayyebi A, Delavar MR, Yazdanpanah MJ (2009) Urban expansion simulation using geospatial information system and artificial neural networks. *Int J Environ Res* 3(4, autumn 2009):493–502. ISSN 1735-6865

- Pontius RG Jr (2000) Quantification error versus location error in the comparison of categorical maps. *Photogram Eng Remote Sens* 66(8):1011–1016
- Pontius RG Jr, Schneider L (2001) Land-use change model validation by a ROC method. *Agric Ecosyst Environ* 85:239–248
- Pontius RG Jr, (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogram Eng Remote Sens* 68:1041–1049
- Pontius RG Jr, Huffaker D, Denman K (2004a) Useful techniques of validation for spatially explicit land-change models. *Ecol Model* 179(4):445–461
- Pontius RG Jr, Shusas E, McEachern M (2004b) Detecting important categorical land changes while accounting for persistence. *Agric Ecosyst Environ* 101(2–3):251–268
- Pontius RG Jr, Spencer J (2005) Uncertainty in extrapolations of predictive land change models. *Environ Plan B* 32:211–230
- Schneider L, Pontius RG (2001) Modeling land use change in the Ipswich watershed, Massachusetts, USA. *Agric Ecosyst Environ* 85:83–94
- Serneels S, Lambin EF (2001) Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. *Agric Ecosyst Environ* 85:65–81
- SPSS INC (2003) SPSS for Windows (Version 12.0). SPSS, Chicago, IL
- Tayyebi A, Delavar MR, Saeedi S, Amini J (2008a) Monitoring the urban expansion by multi-temporal GIS maps. Paper presented at the FIG conference, Stockholm, Sweden
- Tayyebi A, Delavar MR, Saeedi S, Amini J, Alinia H (2008b) Monitoring land use change by multi-temporal LANDSAT remote sensing imagery. In: proceedings of the ISPRS Commission VII, the international society for photogrammetry and remote sensing, Beijing, China, pp 1037–1043
- Tayyebi A, Delavar MR, Pijanowski BC, Yazdanpanah MJ (2009a) Accuracy assessment in urban expansion model. In: Devillers R, Goodchild H (eds) *Spatial data quality, from process to decisions*. Taylor and Francis, CRC Press, Canada, pp 107–115
- Tayyebi A, Delavar MR, Pijanowski BC, Yazdanpanah MJ (2009b) Spatial variability of errors in Urban Expansion Model: Implications for error propagation, *Spatial Data Quality, From Process to Decisions*, Edited by R. Devillers and H. Goodchild, Taylor and Francis, CRC Press, Canada, pp. 134–135
- Veldkamp A, Lambin EF (2001) Predicting land-use change. *Agric Ecosyst Environ* 85(1–3):1–6
- Veldkamp A, Verburg PH (2004) Modelling land use change and environmental impact. *J Environ Manag* 72(1–2):1–3
- Verburg PH, Veldkamp A (2004) Projecting land use transitions at forest fringes in the Philippines at two spatial scales. *Landscape Ecol* 19:77–98
- Wu F, Yeh AG (1997) Changing spatial distribution and determinants of land development in Chinese cities in the transition from a centrally planned economy to a socialist market economy: a case study of Guangzhou. *Urban Stud* 34(11):1851–1880
- Wu J (2004) Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecol* 19:125–138

Chapter 4

Abu Dhabi Island: Analysis of Development and Vegetation Change Using Remote Sensing (1972–2000)

Abdulahakim M. Abdi and Anand Nandipati

Abstract Over the past few decades new cities have appeared around the world in undeveloped areas. And although development has expanded significantly and become bolder and more innovative, the above-average scale at which the countries of the Persian Gulf are growing stands one level above the rest. The United Arab Emirates obtained independence in 1971 with a GDP of 6.5 billion Dirhams (US\$ 1.6 billion); this figure ballooned to 379 billion Dirhams (US\$ 103 billion) in 2004. During this timeframe, the country had undergone tremendous development through petroleum exports and foreign investments. Needless to say, development has permanently changed the country's landscape. The purpose of this paper is to investigate land cover changes in the capital Abu Dhabi and surrounding regions from 1972 to 2000 using Landsat images. Two primary (land and vegetation) and two secondary (shallow and deep water) features were selected as measures of development. Remote sensing and GIS were used to perform the classification and post-classification of images and visualize the results. Results for the two primary features vegetation and land have shown an increase of 3700 and 17% respectively between 1972 and 2000. The creation of new land from the by-products of dredging activities has negative effects on seafloor habitat while the intensified artificial expansion of vegetation impacts groundwater resources, both being direct consequences of rapid development. The application of sustainable methods in development activities is crucial, particularly in this part of the world with very few natural resources other than petroleum and natural gas.

A.M. Abdi (✉)

Instituto Superior de Estatística e Gestão de Informação, Universidade Nova de Lisboa,
Campus de Campolide, 1070-312 Lisboa, Portugal
e-mail: hakim.abdi@gmail.com

4.1 Introduction

Monitoring changes in the environment and the prediction of future circumstances is an important factor in the developing world. The countries in the Persian Gulf are facing tremendous industrialization and their landscape is changing at a tremendous pace (Starbuck and Juanito 2006). Remote sensing plays a significant role in providing geo-information in a spatial format and also in determining and monitoring the overall capacity of the earth (Navalgund et al. 2007).

Prior to independence in December, 1971 the United Arab Emirates (UAE) was a group of loosely associated tribal sheikhdoms called the Trucial States. Petroleum was discovered in the area in 1958 and Abu Dhabi began to export offshore oil in 1962 (Library of Congress 1993). The UAE's proven oil reserves stand at 97.8 billion barrels and as a result of the petroleum wealth the country is getting rapidly industrialized (Alhameli and Alshehhi 2004). The UAE's gross domestic product was Arab Emirates Dirham (AED) 6.5 billion (US\$1.6 billion) in 1972 (El Mallakh 1981), that figure swelled to AED 379 billion (US\$103 billion) in 2004 (UAE-NMC 2008). In 1976, the area under cultivation in the emirate of Abu Dhabi was 28.70 square kilometers (sq.km.) which exploded to 135,000 sq.km. at the end of 1997 (MEAW 2008).

The initial study area for this project was comprised of the original Landsat scene (185 km \times 185 km) covering the central portion of the UAE, however, this approach was abandoned because of the limited time available. Subsequently, the area was reduced to capital, the city-island of Abu Dhabi, and the surrounding region within an average radius of 17 km around the center of the island. This radius was selected because it represents the maximum amount of coastal alteration around the capital.

The current morphology of the island and its surroundings is the result of dredging and land reclamation operations that aim to boost economic productivity. Similarly, afforestation efforts have been extensive in order to beautify the barren landscape of the country (Issa 2008). We found it interesting to compare the change in morphology and the areal extent of vegetation cover from 1972, a year after the emergence of the UAE as an independent country to the year 2000. The rationale for selecting these time periods is that the earliest available satellite imagery for this region dates back to 1972; the year 2000 was chosen because that was when the most recent cloud-free, error-free and cost-free imagery for that particular area was found.

4.2 Data

For this study we used two cloud-free images; a Landsat 1 Multispectral Scanner (MSS) image from November 29th, 1972 and a Landsat 7 Enhanced Thematic Mapper Plus image from August 8th, 2000. Landsat was the chosen satellite because the spatial resolution and the areal extent of the scene cover the region of interest and

Table 4.1 Satellites and sensors used for the study

Satellite	Sensor	Spectral range	Bands used	Pixel Res
L 1	MSS multi-spectral	0.5–1.1 μm	1, 2, 3, 4	60 m
L 7	ETM+ multi-spectral	0.450–1.175 μm	1, 2, 3, 4	30 m

the imagery was freely available. The images were downloaded from the University of Maryland’s Global Land Cover Facility (GLCF) and were orthorectified with a RMS geodetic accuracy of 50 m. The MSS image consists of a pixel size of 60 m and contains four spectral bands; the ETM+ image consists of a pixel size of 30 m and contains eight spectral bands. All MSS bands (1–4) were used while the first four ETM+ bands were used in order to work with a consistent radiometric range on both images. As previously mentioned (see Section 4.1), the August 2000 Landsat 7 image was the most recent available in terms of cost and freedom from error, this is due to the fact that Landsat 7’s Scan Line Corrector (SLC) had been inoperational since May 2003 and this results in a “zigzag” effect on resultant images. About 22% of any scene is lost due to this mishap. The United States Geological Survey (USGS) provides a Gap-filled Systematic Correction with all Landsat imagery since May 2003 but this approach could not be applied for this project due to time constraints (Table 4.1).

4.3 Methodology

The methodology employed in this study involves the classification and post-classification of the satellite imagery which was adapted from the Advanced Training Course on Land Remote Sensing from the European Space Agency (ESA).

The area under analysis consists of low-lying land comprised of tidal mudflats, mangrove forests, reclaimed land made of dredge-spoil plus associated dredged channels, salt-flats, algal mats and an urban landscape. Due to the diversity of landscapes, we chose sets of contiguous pixels as the Spatial Unit of Analysis (SUA). When an image interpretation process is undertaken, one of the key issues in the delineation of discrete areal units on images is the selection of the smallest size area entity to be mapped as a discrete area, the minimum mapping unit (Saura 2002). The chosen Minimum Mapping Unit (MMU) for this study, based on the SUA, is three pixels which correspond to 0.008 sq.km.

4.3.1 Feature Identification and Selection

Due to the limited availability of ancillary data, we have relied solely on aerial imagery for feature selection and identification in the case of the Landsat 1 image. Furthermore, single bands as well as combinations of bands were applied to achieve

maximum possible differentiation. On the other hand, in case of the Landsat 7 image we relied on both visual analysis and local knowledge in identifying and selecting features.

In order to make a clear distinction between the vegetation and other features in both images we employed the Normalized Difference Vegetation Index (Rouse et al. 1973) to both the Landsat images:

$$\text{Normalized Difference Vegetation Index (NDVI)} = \frac{\text{Near Infrared} - \text{Red}}{\text{Near Infrared} + \text{Red}} \quad (4.1)$$

In order to determine the areal extent of afforestation as well as to ascertain the area of reclaimed land, we found it necessary to focus on four information classes; land, vegetation, shallow water and deep water. This would give us the amount of new land created since 1972 as well we the amount of new deep-water channels created as a result of land reclamation efforts and the increase in the amount of vegetation.

4.3.2 Classification

A classification system is designed based on the user's need, spatial resolution of selected remotely sensed data, image-processing and classification algorithms available and time constraints (Lu and Weng 2007). Since this study is not focused on distinguishing between individual land classes apart from vegetation, different land forms such as urban areas, reclaimed islands, salt-flats and low-lying algal mats were aggregated into one region of interest, land. The signatures generated from the training samples are then used to train the classifier to classify the spectral data into a thematic map (Lu and Weng 2007). We used maximum likelihood, a parametric classifier. Maximum likelihood functions on the basis of computation from the records of the classes set from training samples and each pixel is assigned to the class with the highest probability based on training statistics. Polygons of training samples were collected from both images because of the wide generality of the desired classes (Table 4.2).

In the resultant images, there was a distinct "salt and pepper" effect, hence post-classification processing was necessary. For the first stage, a Sieve method was

Table 4.2 Samples collected for Landsat images (1972 and 2000)

1972 Land class	Polygons/pixels	2000 Land class	Polygons/pixels
Deep water	90/7387	Deep water	185/28,863
Shallow water	67/4204	Shallow water	124/25,738
Land	116/21,084	Land	326/110,442
Vegetation	10/39	Vegetation	139/14,458

employed for the purpose of removing the isolated pixels in the classified image. A minimum threshold of three pixels with eight neighbours were selected. Any pixels that did not meet the required parameters were deleted and replaced by unclassified values. Hence, another post-classification method was necessary to merge these unclassified pixels. Majority Analysis (Stuckens et al. 2000) was performed to convert the unclassified pixels within a single class into that class. In this method, a kernel size of 5×5 was specified as well as a center pixel weight of value three, which was replaced based on the values of the kernel of the majority class. After a series of tests with other post-classification methods, this was considered to be a good choice (particularly for a classification with a persistent “salt and pepper” problem) as it is able to eliminate the noise in the classified image. Therefore, after generalization, a smoother image for both the MSS and ETM+ images was produced.

4.4 Results

Determining the accuracy of the classification was performed using a Confusion Matrix. Ground-truth region of interest (GTROI) samples were collected from each image and were assigned the desired classes. These were then entered into the Confusion Matrix and accuracy was assessed against the classified images. As previously mentioned (see Section 4.3.2) the only available ancillary data for 1972 came from aerial imagery and which were also used in the accuracy assessment. The Overall Accuracy of the Landsat 1 image was 94.2% with a Kappa Coefficient (Cohen 1960) of 0.97. The Landsat 7 image produced an Overall Accuracy of 94.07% and a Kappa Coefficient of 0.99 (Table 4.3).

The outcome of the classification displayed a significant increase in the primary features within the temporal range of the study; the final classified images in vector format are presented in Figs. 4.1 and 4.2. Between 1972 and 2000 new land areas were created in the shallow and deep water which amounts to approximately 16% of the total land area. The areal extent of vegetation increased considerably by 3700% during the same period from 3.38 to 127.92 sq.km. These figures were anticipated since the country’s population increased by nearly 10-fold from 309,000 to over 2,776,000 in the 28-year period between 1972 and 2000 (Oxford Business Group 2008; UAE Ministry of Information and Culture 2004) (Figs. 4.3, 4.4, and 4.5).

Even though the study area had undergone tremendous development, some of the original land cover had been retained, for example 86.30% of the original 1972 vegetation cover (coastal mangroves) was retained, which amounts to less than 3% of total vegetation in 2000. Similarly, all the other features experienced a retention percentage of above 80% except shallow water, which had retention of 51.82% of the original areal extent. This is most probably due to the dredging activities that created new land and new deep water channels around Abu Dhabi Island (Table 4.4 and 4.5).

Table 4.3 Accuracy Assessment of the Landsat 1 and 7 images

Percentage accuracy of the Landsat 1 image classification					
Class	Land (%)	Vegetation (%)	Shallow water (%)	Deep water (%)	Total (%)
Land	100.00	31.22	0.00	0.08	29.47
Vegetation	0.00	68.25	0.00	0.00	2.41
Shallow water	0.00	0.53	76.99	0.04	15.57
Deep water	0.00	0.00	23.01	99.88	52.55
Total %	100.00	100.00	100.00	100.00	100.00

Calculated producer's and user's accuracies for the Landsat 1 image classification				
Class	Prod. Acc. %	User Acc. %	Prod. Acc.	User Acc.
Land	100.00	96.13	1514/1514	1514/1575
Land	100.00	96.13	1514/1514	1514/1575
Vegetation	68.25	100.00	129/189	129/129
Shallow water	76.99	99.76	830/1078	830/832
Deep water	99.88	91.17	2561/2564	2561/2809

Percentage accuracy of the Landsat 7 image classification					
Class	Land (%)	Vegetation (%)	Shallow water (%)	Deep water (%)	Total (%)
Land	93.25	15.99	0.00	0.00	33.24
Vegetation	6.75	84.01	0.00	0.00	14.04
Shallow water	0.00	0.00	93.80	0.00	12.65
Deep water	0.00	0.00	6.20	100.00	40.06
Total %	100.00	100.00	100.00	100.00	100.00

Calculated producer's and user's accuracies for the Landsat 7 image classification				
Class	Prod. Acc. %	User Acc. %	Prod. Acc.	User Acc.
Land	93.25	93.25	1519/1629	1519/1629
Vegetation	84.01	84.01	578/688	578/688
Shallow water	93.80	100.00	620/661	620/620
Deep water	100.00	97.91	1922/1922	1922/1963

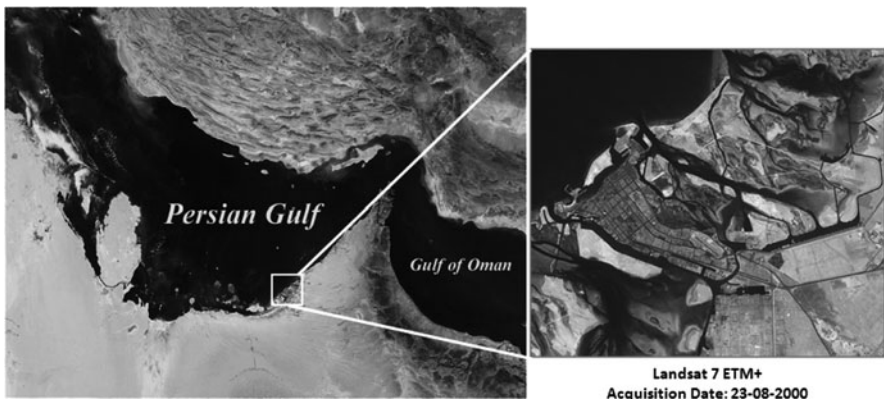


Fig. 4.1 Overview of the study area

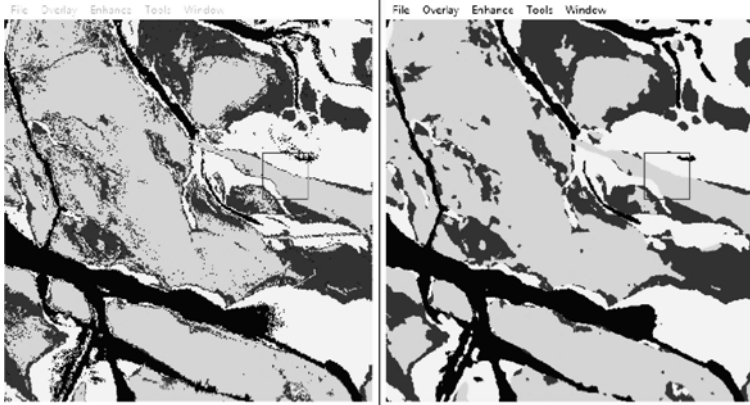


Fig. 4.2 Displays the 2000 classified image before (*left*) and after performing majority analysis

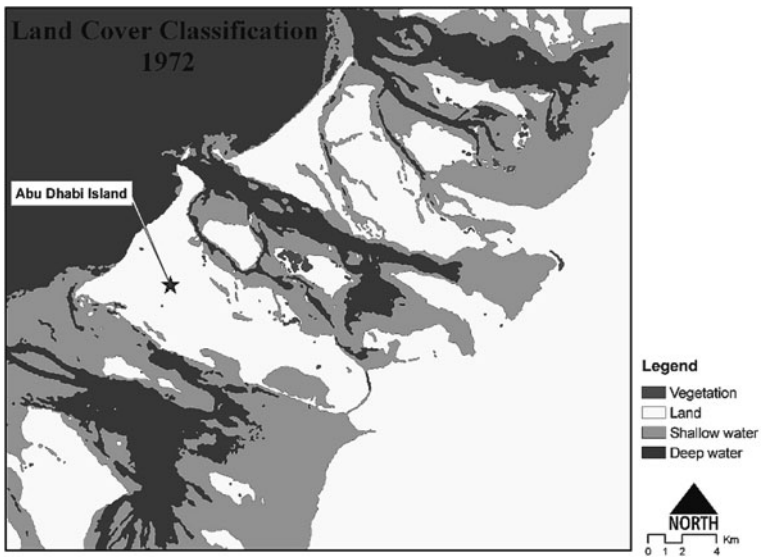


Fig. 4.3 Land cover map of Abu Dhabi in 1972

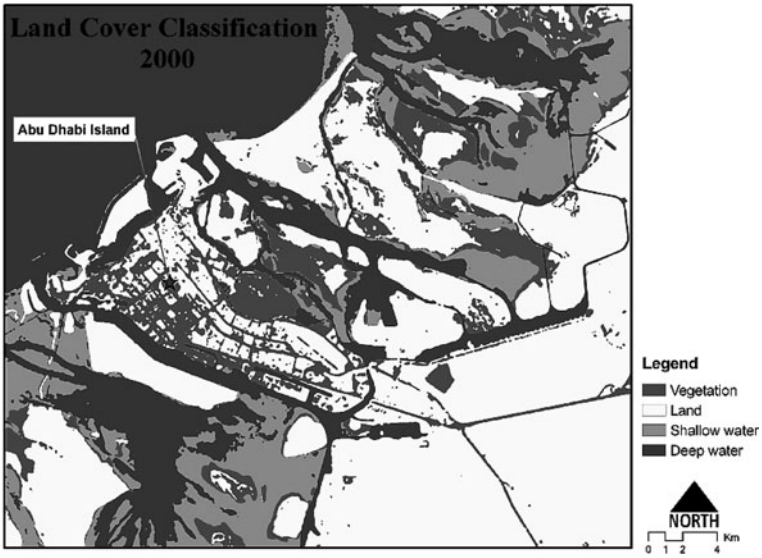


Fig. 4.4 Land cover map of Abu Dhabi in 2000

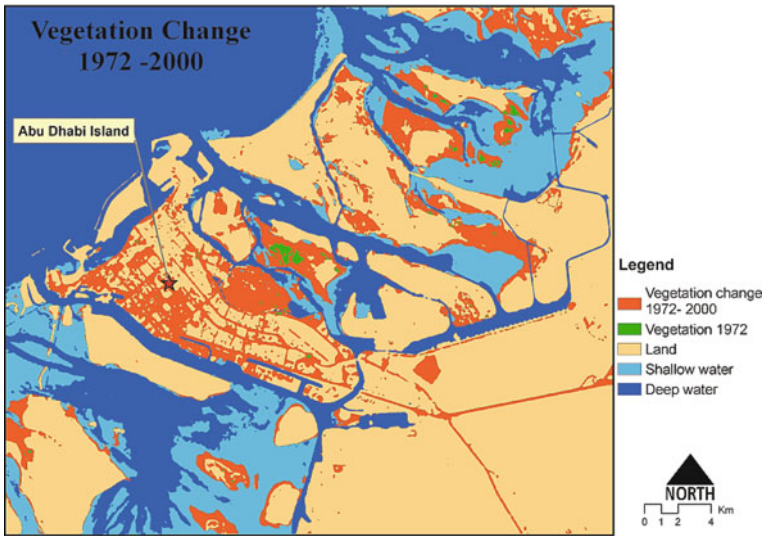


Fig. 4.5 Change in vegetation in Abu Dhabi Island between 1972 and 2000

Table 4.4 Change in different land covers between 1972 and 2000

Change of different land covers to vegetation between 1972 and 2000			
1972 Land cover	2000 Land cover	Areal change (sq.km.)	Percentage based on 2000 land cover area
Vegetation	Land	0.31	0.06
Shallow water	Land	60.62	12.55
Deep water	Land	17.43	3.61
Land	Vegetation	75.51	59.07
Shallow water	Vegetation	48.18	37.69
Deep water	Vegetation	1.22	0.96
Change of different land covers to shallow water between 1972 and 2000			
1972 Land cover	2000 Land cover	Areal change (sq.km.)	Percentage based on 2000 land cover area
Land	Shallow water	2.93	1.52
Vegetation	Shallow water	0.03	0.01
Deep water	Shallow water	32.60	16.93
Change of different land covers to deep water between 1972 and 2000			
1972 Land cover	2000 Land cover	Areal change (sq.km.)	Percentage based on 2000 land cover area
Land	Deep water	17.21	4.93
Vegetation	Deep water	0.12	0.03
Shallow water	Deep water	37.04	10.60

Table 4.5 Amount and percentage of land cover area that was retained between 1972 and 2000

1972 Land cover	1972 Area (sq.km.)	2000 Land cover	2000 Area (sq.km.)	Areal retention (sq.km.)	Retention percentage based on 1972 land cover area
Land	500.42	Land	483.13	404.77	80.89
Vegetation	3.38	Vegetation	127.83	2.92	86.39
Shallow water	302.76	Shallow water	192.48	156.92	51.82
Deep water	344.24	Deep water	347.36	292.99	85.11

4.5 Discussion

The countries of the Persian Gulf have simultaneously outpaced their neighbors in the development and modernization since the discovery of petroleum. However, this rapid development was not monitored until recently when the science of remote sensing began to be taken up by academia and researchers in the region. We have demonstrated that remote sensing can be used to measure the amount of change over the course of several years in rapidly changing landscapes. This method can

be applied to other cities in the region, and indeed worldwide, that have undergone similar levels of development. Although the scope here was to see the change in land and vegetation over 28 years, future work of interest may be the evaluation of how the urban fabric and impervious surfaces of Abu Dhabi have increased since independence and their contribution to the Urban Heat Island effect (Jones et al. 1990).

4.6 Conclusions

The area in and around the capital city of Abu Dhabi was subject to intensive development efforts that permanently altered the landscape. These drastic morphological changes have been attributed to the surge in petroleum wealth. This study has successfully produced a land cover map of the selected study area through the application of a supervised classification algorithm. This study demonstrated that remote sensing can be used in the monitoring of landscapes that quickly change in relatively small temporal scales.

Albeit the scope of this study is limited, it does provide valuable insight into the extent to which the UAE is changing. More research on development and its impacts on landscapes are needed, particularly in rapidly developing nations such as the UAE, is required and remote sensing provides an invaluable tool.

Acknowledgments This work has been supported by the European Commission, Erasmus Mundus Programme, International M.Sc. in Geospatial Technologies, project no. 2007-0064. We would like to thank Dr. Mario Caetano of Instituto Superior de Estatística e Gestão de Informação, Universidade Nova de Lisboa and Instituto Geográfico Português for his invaluable support and advice in carrying out this study.

References

- Alhameili S, Alshehhi M (2004) Images are an outstanding evidence of rapid development “A perfect example from United Arab Emirates (UAE)”. ISPRS XXth Congress, Istanbul, Turkey, p 505
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20(1):37–46
- El Mallakh R (1981) The economic development of the United Arab Emirates. Croom Helm, London
- Issa S (2008) GIS approach for assessment of human footprint and land management: an example in the United Arab Emirates. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 591–595
- Jones PD, Groisman PY, Coughlan M, Plummer N, Wang WC, Karl TR (1990) Assessment of urbanization effects in time series of surface air temperature over land. *Nature* 347:169–172
- Library of Congress, Federal Research Division (1993) Retrieved on 12 December 2008. United Arab Emirates: a country study, <http://hdl.loc.gov/loc/gdc/cntrystd.ae>
- Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance. *Int J Remote Sens* 28(5):823–870
- MEAW, Department of Agriculture (2008) Retrieved on 12 December 2008, from Ministry of Environment and Water, <http://www.moew.gov.ae/AR/Pages/default.aspx>
- Navalgund R, Jayaraman V, Roy P (2007) Remote sensing applications: an overview. *Curr Sci* 93:1747–1766

- Oxford Business Group (2008) The report: Abu Dhabi 2008. Oxford Business Group, London
- Rouse JW, Haas RH, Schell JA, Deering DW (1973) Monitoring vegetation systems in the Great Plains with Earth Resources Technology Satellite. In: Proceedings of 3rd Earth Resources Technology Satellite Symposium, Washington, DC, pp 309–317
- Saura S, (2002) Effects of minimum mapping unit on land cover data spatial configuration and composition. *Int J Remote Sens* (23)22:4853–4880
- Starbuck MJ, Juanito T (2006) Monitoring vegetation change in Abu Dhabi Emirate from 1996 to 2000 and 2004 using Landsat satellite imagery. In: Proceedings of 2nd International Kuwait Conference, Kuwait City, pp 817–831
- Stuckens J, Coppin PR, Bauer ME (2000) Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sens Environ* 71:282–296
- UAE Ministry of Information and Culture (2004). UAE yearbook 2004. Trident Press, London
- UAE-NMC, Gross Domestic Product (2008). Retrieved on 12 December 2008, from United Arab Emirates National Media Council, <http://uaeinteract.com/news/default.asp?ID=242>

Chapter 5

Remote Sensing of Coastal Water Quality in the Baltic Sea Using MERIS

Andreas Neumann, Harald Krawczyk, and Stefan Riha

Abstract Remote sensing of water quality parameters using spaceborne imaging spectrometers has become a widely spread technology during the last years. Especially MERIS on ESA's ENVISAT with its 250 m spatial resolution provides new opportunities for applications in coastal and inland waters. To make use of the information content of spectrally highly resolved data new retrieval algorithms had to be developed based on bio-optical and radiative transfer modeling and inversion by the use of principal component analysis. The paper will review the approaches and algorithms developed at DLR for research and operational applications to assess and monitor water quality in case-2 waters (coastal and inland waters). Special focus will be on yellow substance dominated waters and exceptional algal blooms in the Baltic Sea. In addition to the methodological aspect the processing system and infrastructure to provide daily water quality services in the frames of GMES will be introduced.

5.1 Introduction

During the last years spaceborne imaging spectrometers have been providing higher spectral resolution data for ocean colour remote sensing in the visible-near-infrared spectral range (400–1000 nm). To some extent MOS-IRS (Zimmermann 2000), launched in 1996, was a pacemaker in this respect and in-between was followed by several instruments both of experimental and pre-operational status (e.g. CHRIS, MERIS). Where on one hand this new class of instruments allows to access, distinguish and quantify more parameters by remote sensing with increased accuracy, the spectral dimensionality of the data require new retrieval techniques accounting for the multivariate nature of the data (i.e. several parameters may vary independently and influence the spectral signature measured by the instrument)

A. Neumann (✉)

Remote Sensing Technology Institute, German Aerospace Center DLR,
Rutherfordstrasse 2, 2489 Berlin, Germany
e-mail: andreas.neumann@dlr.de

and allowing for regional and/or seasonal optimization of the inversion algorithms. In the last years this led to the development of bio-optical and radiative transfer model based techniques fulfilling these requirements. Examples are Principal Component Inversion PCI (Krawczyk et al. 1993), Neural Network algorithms NN (Doerffer and Schiller 1998; Niang et al. 2003; Zhang et al. 2002) or Spectral Matching techniques SM (Heege et al. 2005; Morel and Bélanger 2006; Van Der Woerd and Pasterkamp 2008). Since these techniques follow a consistent physical model and information theory based approach which is different to “classical” empirical algorithms still often used in remote sensing we will discuss in the following the physical and mathematical basis and how in particular PCI can be used to implement dedicated case-2 inversion algorithms. The algorithms developed were implemented in an operational processing scheme used for water quality monitoring in the Baltic Sea in the frames of the European Global Monitoring for Environment and Security (GMES) programme for German users. Some examples will illustrate this operational application.

5.2 Characteristics of Optically Complex Waters

This paragraph will summarise the main points on this subject, a more extensive discussion can be found in Sathyendranath (2000). Open ocean waters, which are usually characterised by phytoplankton being the main varying water constituent, can be considered as a single-variable system from the remote sensing perspective. Chlorophyll as the dominating optically active substance is determining the “colour” (i.e. reflectance spectrum) of the water body. For satellite measurements, i.e. top of the atmosphere, additional variability in the measurement is caused by atmospheric variability (e.g. aerosols, gaseous absorbers). In most cases for the atmospheric correction the atmosphere is described by Rayleigh scattering, aerosol content and a parameter related to the size distribution of aerosols. Thus, for the inversion of satellite ocean colour measurements in total three main variables seem sufficient for open ocean waters (we neglect here additional variables like absorbing gases assuming a proper positioning of the instrument’s spectral channels to minimise their influence).

Because of the strong light absorption by water at wavelengths above 700 nm and the comparably low phytoplankton concentrations usually the “black water condition” for atmospheric correction firstly introduced by Gordon (1978) can be applied and band ratios using two or three channels in the visible range can be used to retrieve chlorophyll concentrations.

The situation is different in inland and coastal waters or (semi-) closed basins, such as the Baltic Sea or Black Sea, due to one or several of the following reasons:

- significantly higher phytoplankton concentrations occur, this may result in a non-negligible signal from the water body at wavelengths >700 nm
- occurrence of several different and statistically independent optically active constituents in the water body will lead to interpretation failures of “one-dimensional” retrieval algorithms

- large concentrations of inorganic scatterers in the water (in particular suspended sediment) cause a significant signal remitted at wavelengths > 700 nm and violate the “black water condition” for atmospheric correction
- land influence may cause variations in aerosol composition (i.e. mixtures of maritime and land aerosols) and may require determination of aerosol properties from the measurement to be used in atmospheric corrections.

For the optically active water constituents we find the following situation:

- the main groups of constituents considered are phytoplankton, dead organic material (Seston), suspended inorganic matter (sediment) and dissolved organic matter (Gelbstoff, DOM)
- the composition of phytoplankton species, and thus the optical properties and resulting colour characteristics, may vary significantly with region or season
- due to different origin the colour characteristics of suspended sediments may vary significantly with place and time
- extremely high concentrations of one or more constituents (e.g. in river plumes or algae blooms) may optically mask other components
- in general, the different components in the water do not covary due to their different biological or physical nature, different origin and different spread mechanisms.

In result of this all specific requirements become valid for retrieval algorithms to be applied to optically complex (case-2) waters:

- algorithms are needed capable to account for simultaneous occurrence of several, non-covarying water constituents
- algorithms and models to be applied need to account for specific optical properties of different water constituents to be able to discriminate between them
- a retrieval scheme is needed which allows to account for changing, regional specific optical properties of the water body and
- advanced atmospheric correction algorithms are needed to account for residual water reflectance in the near infrared as well as for varying aerosol properties.

Thus, by dealing with optically complex waters, typically found in coastal and inland waters, we are facing a multivariate mixture of variables in the measurement. This, finally, leads to the demand for more spectral bands and better spectral resolution in the measurement and for adequate, multivariate inversion techniques to retrieve the geo-physical parameters.

5.3 Physics of the Problem

The “Colour” of the water body which is the basis for remote sensing of water constituents is caused by the scattering and absorption properties of pure water and the individual constituents respectively. These properties are described

by the corresponding scattering and absorption spectra, or, if normalised to concentrations, the specific spectral scattering and absorption coefficients (inherent optical properties, IOPs). This means, what we “see” by measuring the reflectance spectra, are not the concentrations of water constituents, but the result of several superimposed scattering and absorption processes forming the resulting spectrum.

This understanding is essential for developing retrieval algorithms because it means, from the physical point of view, that primarily we are measuring scattering and absorption properties by remote sensing. These are then implicitly or explicitly linked to concentrations of the single water constituents, implicitly in the case of empirical algorithms correlating reflectances with in situ concentrations, explicitly through the specific IOPs in the case of model-based inversion techniques.

Another important conclusion follows from this physical background: since the scattering and absorption spectra of the single components are smooth, i.e. without strong spectral features (except special effects like Fluorescence which are not considered here) and superimposing in the same spectral range it is almost impossible to retrieve one single parameter without accounting for the occurrence and variation of the others. This, finally, is the reason for the necessity of a larger number of spectral bands, higher spectral resolution and extended spectral range for remote sensing of optically complex case-2 waters.

5.4 Mathematical View

To better understand the function and problems of the retrieval (or inversion) algorithms it is worth to look at the problem from a mathematical and information theory perspective. The spectra measured for case-2 waters remote sensing are a multivariate data set, i.e. variations in the measurement are caused by multiple parameters which are not necessarily co-varying with each other. The spectral measurements in each of the spectral channels of an instrument are statistically not independent due to the smooth and “broadband” optical properties of the individual constituents in the water. It is therefore hardly possible to correlate single bands to the variation of an individual constituent or one optical parameter. In fact, each available channel contains (part of) information on each of the water constituents. Thus, multivariate inversion techniques are the adequate tool to build retrieval algorithms for optically complex waters. This type of inversion technique is capable to use as much as possible spectral information (i.e. use all available spectral bands) to estimate an individual parameter while at the same time accounting for other variables influencing the spectrum.

How this approach translates into concrete algorithms? Figure 5.1 shows the basic scheme for the inversion of a multivariate spectral measurement. In principle what is needed is a multidimensional matrix of weighting coefficients which “maps” (transforms) the spectral reflectance values to the desired geo-physical values. The determination of the coefficients can hardly be done using empirical correlations, since the number of possible parameter combinations would require an unrealistically large set of measured spectra plus corresponding in situ data.

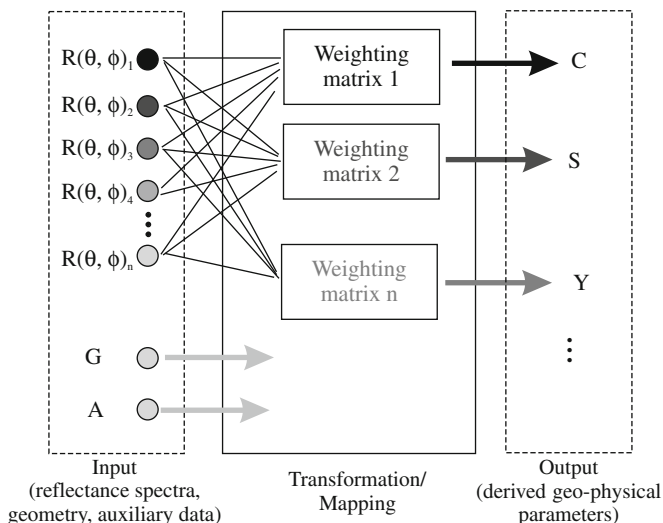


Fig. 5.1 Scheme of a multivariate inversion of spectral (remote sensing) measurements

Radiative transfer modelling using the specific IOPs of the relevant water constituents and atmospheric parameters provides us with a possible solution. The IOPs, of course, must be known from in situ and laboratory measurements, atmospheric parameters from corresponding measurements or models. The modelling then allows to generate a synthetic data set containing all possible combinations of geo-physical parameters and the corresponding reflectance spectra including observation and sun geometries for a given sensor/satellite. This data set represents the expected variety of measurements of the sensor. This “training” data set may then be used to compute the weighting coefficients for the desired inversion. Numerical techniques have to be applied because a direct (analytical) inversion of the model is not possible.

There exist a number of mathematical methods to solve the inverse task. Examples are: factor analysis, principal component inversion, neural networks or spectral matching techniques. Due to different mathematical treatment in detail, the resulting inversion schemes or algorithms allow different optimisation strategies and show different behaviour in the case of non-linearities or large variation ranges of the parameters. But the basic approach for this class of algorithms is identical, as described above. Figure 5.2 illustrates the scheme.

5.5 Atmospheric Correction

As outlined above, the classic approach for atmospheric correction fails in the case of high concentrations of scattering constituents in the water (e.g. high sediment load). Atmospheric correction in these cases needs to account for residual signal

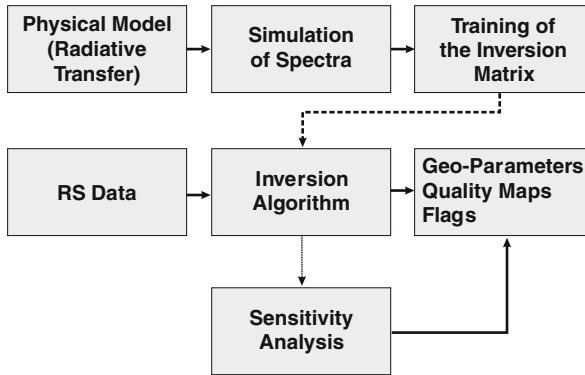


Fig. 5.2 Scheme of training an inverse algorithm

from the water body also in the near infrared, i.e. above 700 nm. This, as one solution, leads to more complex or iterative atmospheric correction schemes, which in addition to aerosol properties also estimate residual in-water scattering. Following this way to the end, an even more integrated view on the problem can be developed: Applying an atmospheric correction to the data separately, i.e. computing water leaving radiances or reflectances from the top-of-atmosphere (ToA) measurements does not improve the information content in the data with respect to water constituents. It only removes variables related to the atmosphere from the original data set and hence reduces the dimensionality for the inversion. But all variation due to changing water constituents must be resolved in the ToA measurement already, otherwise it may not be seen in atmospherically corrected data. Taking this into account and also reminding that atmospheric correction for case-2 waters must account for residual signal from the water body, it should be possible to couple the retrieval of aerosol parameters and water constituents into an integral inversion or coupled inversion scheme. For this case the training data set for multivariate inversion is simulated for TOA reflectances by including variations of aerosol-related parameters. To derive the inversion schemes the identical mathematical apparatus is used as for water leaving reflectances, but a larger dimensionality has to be accounted for.

Such an approach we tried first for the imaging spectrometer MOS-IRS flying on board of the Indian satellite IRS-P3. The mission was a joint effort of DLR and Indian Space Research Organisation ISRO, where the instrument was developed and built by DLR and ISRO provided the satellite and the launch and operated the satellite. Being planned as a proof-of-concept experiment it was flying successfully for 8 years from 1996 to 2004 (Zimmermann and Neumann 2000). MOS-IRS was the first imaging spectrometer in space for Earth observation and provided measurements in 18 spectral channels between 408 nm and 1.6 μm . Because of very similar instrument characteristics it was used as a precursor for MERIS, allowing for experimental verification of algorithms developed for MERIS on ENVISAT, which was launched in 2003. The Principal Component Inversion technique considered here was firstly developed and tested using MOS-IRS (Krawczyk et al.

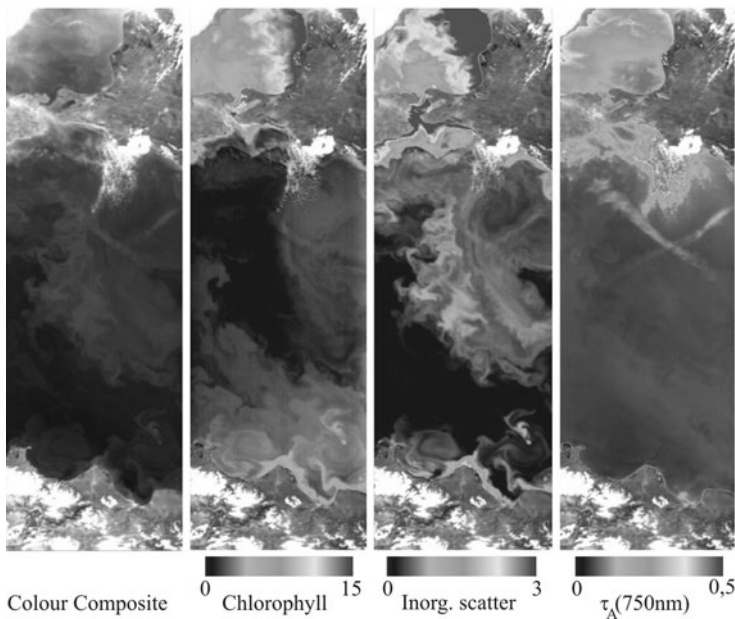


Fig. 5.3 Example for one-step inversion (MOS-IRS, central Black Sea)

1998) data and then extended to be applied to MERIS. Extensive radiative transfer simulations were used to derive and optimise the inversion procedure, the results obtained from MOS-IRS confirmed the simulations and showed very good performance, in particular for case-2 waters. Figure 5.3 shows an impressive example for a very complex and dynamic scene over the Black Sea: where the composite image on the left shows the mixture of all components (water and atmosphere) contributing to the signal measured at the satellite the inversion results illustrate that constituents showing absorption plus scattering (Chlorophyll) are well discriminated from constituents with dominant scattering (inorganic sediment) and the atmospheric component represented by the aerosol. However, at large value of scattering near the coasts saturations and masking effects occur. The methodology is currently tested for MERIS data where the larger variation in viewing geometry due to the large swath of 1500 km makes the problem more complex (MOS-IRS had a 200 km swath). A similar approach using neural networks for inversion has been implemented for MERIS by Schroeder et al. (2007).

5.6 Principal Component Inversion

Here we can only give a short introduction to Principal Component Inversion, for more details are referred to Krawczyk et al. (1993, 1998). The forward model, which computes the reflectance spectra R from the given set of geo-physical parameters p ,

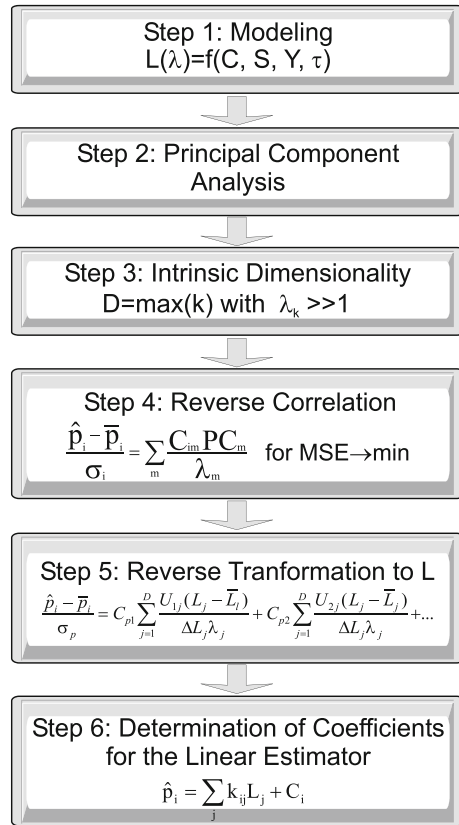
is nonlinear and consists of a radiative transfer code and the necessary inherent optical properties of the components, as there are the wavelength-dependent specific absorption and scattering coefficients and the scattering phase functions. The inverse mapping $R \rightarrow p$ is the desired interpretation algorithm. For PCI the main idea is to estimate the concentration as a linear function of the measured spectral reflectances. But this assumption initially contradicts the nonlinear character of the direct model. Two steps were taken to improve the situation. Firstly instead of the direct parameter p a semi logarithmic representation $q = p + 0.1 \log(p)$ is used, secondly, the covered variability range for each parameter p is divided into sub-ranges, where the linear assumption is much better justified (quasi-linear approach). The estimator then can be formulated as:

$$q = p + 0.1 * \log(p) = \sum_{j=1}^n k_{ij}^{(r)*} R_j + b_j \quad (5.1)$$

where k_{ij} and b_j are the inversion coefficients to be determined for each sub-range r from the model, n is the number of spectral channels. The estimation shall be optimal in a global sense, i.e. the RMS error of the entire dataset should be minimized. This is a difference to so called direct model inversion methods (e.g. spectral matching), which try to minimize every individual spectrum by finding the optimal concentration set for a given model and single spectrum. Here a local linear regression technique is used to compute the needed coefficients. In this task the inversion of the regression matrix of reflectances is a necessary step. Due to the high spectral correlations this can lead to massive numerical problems. Therefore a regularisation method must be applied, to overcome this ill-posed problem. One also must take into account the radiometric resolution and noise in the measurement. As an optimal information extraction and noise suppression tool, the principal component analysis (PCA) was chosen. Figure 5.4 illustrates the steps to compute the coefficients.

The first step is the simulation of a large set of reflectances R or radiances L as explained above. Then a principal component analysis is applied to this data set. The eigenvalues λ determine the intrinsic dimensionality, i.e. the principal components corresponding to the highest eigenvalues contain the main and useful part of information and the lower eigenvalues corresponding components contain the measurement noise. These lower principal components are omitted for further calculations to increase noise stability of the developed inversion algorithm. Next the correlation between geophysical parameters and higher principal components is established (step 4). Since the principle components are orthogonal it can be easily done. But his formula can not yet be used for a general interpretation, because the result of PCA strongly depends on the statistical (covariance) properties of the initial data set. The data set in a natural scene will never be expected the same, as that used for the simulation. Therefore this formula must be generalized. This can be done by back-transforming the principal components to radiances using the eigenvectors. Finally one gets a regression formula between parameters and radiances (steps 5

Fig. 5.4 Steps of the principal component inversion



and 6). Comparing with a Neural Network approach one could assess, both methods are performing a model inversion, minimizing the global interpretation error and differing mainly in the method of “training” the interpretation coefficient sets. Neural nets are often using backpropagation techniques, PCI uses principal component transformation as an optimized error-noise suppressing filter. One advantage of the PCI is the additional information about correlations between the parameters and principal components, which allows a direct estimation of the interpretation potential of the investigated data set. Concerning the piecewise linearization of the data set during interpretation one has to choose the appropriate of the pre-calculated coefficient sets. The problem is solved by trying all sets and testing the sub-range conditions under which they were calculated. In the case of impossibility to find a solution an invalid flag is raised.

Similar basics as described here for the Principal Component Inversion are used in the Minimum Noise Fraction Transformation (MNF, e.g. Chen 2000). However, there it is not the goal to build a physically based inversion to derive geo-physical parameters directly from remote sensing measurements but to find

optimal algorithms to reduce dimensionality and noise in the data. Since in the described above Principal Component Inversion we use Eigenvectors normalised to expected measurement noise the optimization effect is similar to the approach of the MNF.

5.7 Monitoring of Water Quality

Monitoring of (coastal) water quality by the use of remotely sensed data has become an important task during the last years (Kratzer et al. 2008). In Europe the Global Monitoring for Environment and Security (GMES, <http://www.gmes.info/>) programme by ESA and the European Commission has put a significant effort into the development of infrastructure and services providing data and information for coastal environmental management and the monitoring of environmental directives. One project dealing in particular with water quality information and surveillance of critical algal blooms in European waters is the “Marine and Coastal Environmental Services” (MarCoast) funded by ESA (<http://gmes-marcoast.com/>). At DLR in the frames of this project an operational processing chain for MERIS data has been set up and integrated in the automated environment at the German Remote Sensing Data Center. It serves users in Germany (governmental and public institutions on national and state levels) with information products on a daily basis. The products are available in the form of images as well as scientific data formats, optional GIS-compatible formats are also available. All products are provided as daily maps and 10-day, monthly and seasonal averages.

Water quality:

- Water constituents (Chlorophyll, suspended sediment, Gelbstoff)
- RGB colour composites
- Sea surface temperature
- Water transparency

Algal bloom monitoring:

- Bloom strength indicator map
- Bloom location and extend (text)

The services are running now for three years and will be continued on a regular basis. User access is realised via an ftp-server, in addition all data are archived in DLR’s multission long term archive system. Figures 5.5 and 5.6 shows examples for the set of information products. The results in terms of geophysical parameters have been validated against in situ measurements from regular monitoring stations operated by state authorities of Mecklenburg-Vorpommern (German state along the Baltic coast). Chlorophyll and sediment show good agreement except very near to

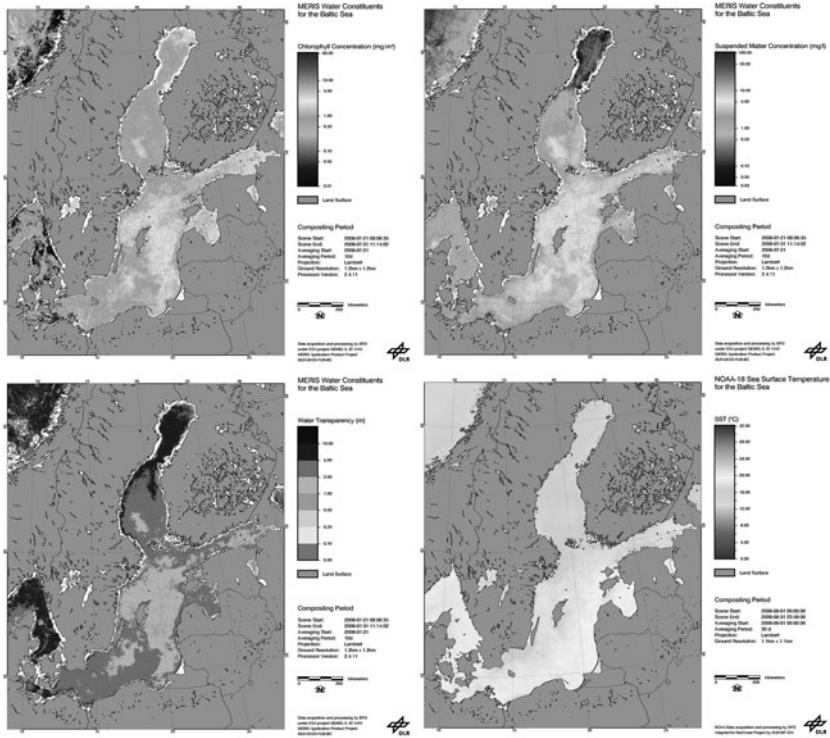


Fig. 5.5 Operational water quality products of the Baltic Sea derived from MERIS (top left to bottom right: Chlorophyll, suspended matter, water transparency, sea surface temperature)

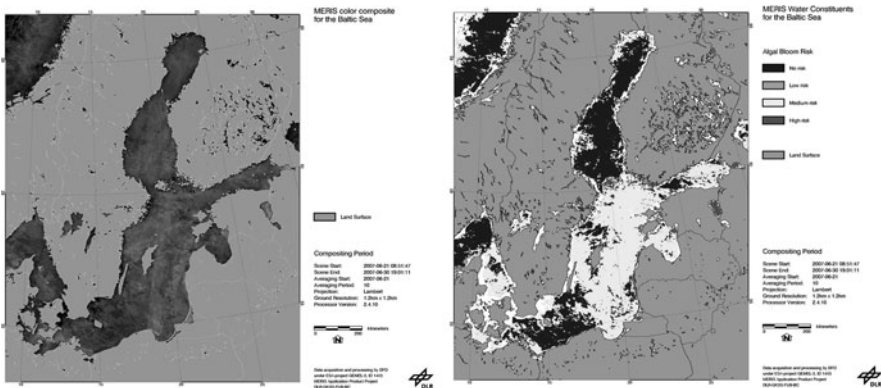


Fig. 5.6 Products for the monitoring of algal blooms (left: color composite, right: bloom strength indicator)

the coast where bottom and adjacency effects cause problems. The accuracy of yellow substance, by its physical nature the weakest parameter for retrieval, is currently not satisfactory. One cause for this may be inappropriate bio-optical model, is this is currently investigated.

5.8 Conclusions

The physical complexity of coastal waters result in a challenge to inversion algorithms applied to data over such waters. Model-based, multivariate inversion techniques provide the adequate tool to solve the task. During the past years this technique using a variety of mathematical methods and implementations has become mature and finds a wide range of applications for environmental monitoring. Hyperspectral instruments which will become more and more available in the upcoming years provide measurements with a dimensionality of an order of magnitude larger than currently operational satellites. Also in this case multivariate algorithms are the tool to cope with this huge information content.

References

- Chen C-M (2000) Comparison of principal component analysis and minimum noise fraction transformation for reducing the dimensionality of hyperspectral imagery. *J Geophys Res* 33:163–178
- Doerffer R, Schiller H (1998) Determination of case-2 water constituents using radiative transfer simulation and its inversion by Neural Networks. In: *Proceedings of the Ocean Optics XIV, Kailua-Kona, Hawaii, USA*
- Gordon HR (1978) Removal of atmospheric effects from satellite imagery of the oceans. *J Appl Optics* 17(10):1631–1636.
- Heege T, Miksa S, Kisselev V (2005) Physically based methodology for coupled retrieval of aerosol and water composition from hyperspectral remote sensing data. In: *Proceedings of the 3rd International Conference "Current Problems in Optics of Natural Waters"*, St. Petersburg, Russia, pp 205–209.
- Kratzer S, Brockmann C, Moore G (2008) Using MERIS full resolution data to monitor coastal waters – a case study from Himmerfjärden, a fjord-like bay in the northwestern Baltic. *Remote Sens Environ* 112(5):2284–2300
- Krawczyk H, Neumann A, Hetscher M, Zimmermann G, Walzel T (1998) Principal component analysis a tool for interpretation of multispectral remote sensing data. In: *Proceedings of the Ocean Optics XIV, Kailua-Kona, Hawaii, USA*
- Krawczyk H, Neumann A, Walzel T, Zimmermann G (1993) Investigation of interpretation possibilities of spectral high dimensional measurements by means of principal component analysis – a concept for physical interpretation of those measurements. *SPIE Proc* 1938:401–411.
- Morel A, Bélanger S (2006) Improved detection of turbid waters from ocean color sensors information. *Remote Sens Environ* 122(3–4):237–249
- Niang A, Gross L, Thiria S, Badran F, Moulin C (2003) Automatic neural classification of ocean colour reflectance spectra at the top of the atmosphere with introduction of expert knowledge. *Remote Sens Environ* 86(2):257–271
- Sathyendranath S (ed) (2000) *Remote sensing of ocean colour in coastal and other optically-complex waters*. Bedford Institute of Oceanography, Canada. IOCCG Report No. 3

- Schroeder Th, Schaale M, Fischer J (2007) Retrieval of atmospheric and oceanic properties from MERIS measurements: a new Case-2 water processor for BEAM. *Int J Remote Sens* 28(24):5627–5632
- Van Der Woerd HJ, Pasterkamp R (2008) HYDROPT: a fast and flexible method to retrieve chlorophyll-a from multispectral satellite observations of optically complex coastal waters. *Remote Sens Environ* 112(4):1795–1807
- Zhang Y, Pulliainen J, Koponen S, Hallikainen M (2002) Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data. *Remote Sens Environ* 81(2–3):327–336
- Zimmermann G, Neumann A (2000) The imaging spectrometer experiment MOS on IRS-P3 – three years of experience. *J Spacecr Technol* 10(1):1–10.

Chapter 6

Remote Sensing for Water Quality Monitoring in Apalachicola Bay, USA

Wenrui Huang, Shuisen Chen, and Xiaojun Yang

Abstract In this paper, we provide a review on remote sensing of water quality in coastal and estuarine ecosystems. As a case study, we further describe the use of remote sensing to estimate total suspended solids (TSS) and chlorophyll-a in an estuarine ecosystem, Apalachicola Bay, which is located along the northeastern Gulf of Mexico, USA. Based on the remote sensor data and the regression models derived for the Apalachicola Bay, we found that TSS concentrations indicate strong sediment resuspension, which may be induced by the passage of Hurricane Frances in 2004. We also examine the effects of river flow on estuarine chlorophyll-a concentrations by using remote sensor data, showing higher chlorophyll-a concentrations during the high-flow season. This study demonstrates that remote sensing can be used as an effective tool for water-quality monitoring in coastal ecosystems.

6.1 Introduction

Water quality is important for healthy estuarine ecosystems. Estuarine water quality can be affected by many natural and anthropogenic factors. For example, in shallow water estuaries of the Gulf of Mexico, tides, wind, and river flows can easily re-suspend sediments, causing high turbidity and changes in salinity regimes (Huang 2009; Huang et al. 2002; Wang et al. 2009). Among many estuaries along the Gulf of Mexico, Apalachicola Bay is a highly productive barrier-island estuary, which generally produces 90 percent of Florida's commercial oyster harvest and the third largest shrimp catch (Whitfield and Beaumariage 1977). It is located in the panhandle of Florida, USA (Fig. 6.1). Preservation of the ecology

W. Huang (✉)

Department of Civil Engineering, Florida State University, Tallahassee, FL 32310, USA
e-mail: whuang@eng.fsu.edu



Fig. 6.1 Location of the Apalachicola Bay, Florida, USA

in Apalachicola Bay has been recognized as of state, federal, and international importance. The bay has been designated as a National Estuarine Research Reserve, Outstanding Florida Water, State Aquatic Preserve, and International Biosphere Reserve. The importance of freshwater inflow to the estuarine productivity and the aquatic ecosystem has been recognized by researchers and coastal managers for decades (Livingston et al. 2000). The high seafood production in estuarine systems is often associated with sufficient freshwater input. Currently, the Bay is in relatively unaltered state with freshwater inflow as a major controlling factor to the ecosystem (Livingston 1984; Livingston et al. 2000). Oyster mortality due to stenohaline predators (salt-water fish) is a major determinant of oyster productivity in the bay. Low salinity resulting from sufficient freshwater input limits the predators from the oyster beds and maintains the oyster productivity (Livingston et al. 2000). Fresh water carrying nutrients is also a key factor to support the growth of oysters, shrimps, and other species in the aquatic ecosystem. Modification of river flow patterns and reduction of flow rates will affect the circulation and salinity in the bay. As a result, it will affect predator-induced oyster mortality and growth (Livingston 2000).

Apalachicola Bay receives freshwater flows from the Apalachicola, Chattahoochee, and Flint River system (ACF), which drains over 60,000 km² of Georgia, Alabama, and Florida (Livingston 2000). Consequently, human activities such as land use change, development, and upstream water diversion in the ACF basin greatly affect the water quality of Apalachicola Bay. The spatial and temporal patterns of water quality in the Bay reflect the spatial and temporal heterogeneity of these controlling factors. The spatial and temporal variability in chlorophyll-*a* concentration and TSS greatly affects the spatial and temporal variations in oyster production in the Bay (Wang et al. 2009). Hence, detecting and mapping water

quality across the entire Bay are crucial to environmental monitoring and can help identify sensitive and vulnerable areas for detailed study, management, or restoration in the Apalachicola Bay system. Because the distributions of water quality variables in shallow Gulf of Mexico estuaries are spatially and temporally variable, one of the best, synoptic approaches to understand the variations in water quality variables to support resource management practices is through remote sensing. Remote sensor data from the Moderate-Resolution Imaging Spectroradiometer (MODIS) can be effective in detecting water quality in bays and estuaries because of their medium but sufficient spatial resolution, sufficient sensitivity to water quality variables, especially suspended sediment, and nearly daily coverage.

6.2 Water Quality Indicators for Coastal and Estuarine Ecosystems

Water quality parameters such as chlorophyll-a, total suspended solids (TSS), salinity, temperature, and velocity can be useful indicators of ecosystem health for oyster growth. The spatial and temporal patterns of water quality in the Bay reflect the spatial and temporal heterogeneity of these controlling factors. Sediments in shallow waters can impact the physical and ecological environment of a water column through sediment re-suspension and transport (Wang et al. 2008). High TSS concentration can cause significant mortalities of oysters during the earliest phase of larval development, and result in decreased oyster filtration rates (e.g. Dekshenieks et al. 2000; Wang et al. 2008). Field data measurements of sediment concentrations are generally expensive or have limited spatial coverage, especially for a large area of estuarine water (Zawada et al. 2007). The distributions of TSS concentration in the shallow Gulf of Mexico estuaries are spatially and temporally variable, especially during a hurricane. Remote sensing provides an effective approach to monitor TSS concentrations in the large water area to support the monitoring of hurricane-induced sediment re-suspension and resource management practices in Apalachicola Bay.

Chlorophyll-a concentration is an indicator of phytoplankton abundance and biomass of microscopic plants (phytoplankton), such as unicellular algae in coastal and estuarine waters. They are potential indicators of maximum photosynthetic rate. High levels of Chlorophyll concentration often indicate poor water quality and low levels often suggest good conditions. However, elevated chlorophyll-a concentrations are not necessarily a bad thing, unless it is the long-term persistence of elevated levels. An increase in chlorophyll-a indicates potential eutrophication of the system. Consistently high or variable chlorophyll-a concentrations may indicate the occurrence of algal blooms, which can be harmful to other aquatic organisms. Chlorophyll-a is most likely to respond to nutrients and hydrodynamics; it may also respond to changes in freshwater flow regime, water temperature, and aquatic sediments.

6.3 Remote Sensing for Estuarine Water Quality Monitoring

6.3.1 TSS Monitoring

Water quality in estuarine ecosystems is affected by many natural and anthropogenic factors. Water quality parameters such as colored dissolved organic matters (CDOM), chlorophyll-a, total suspended solids (TSS), salinity, temperature, and velocity can be useful indicators of ecosystem health. The spatial and temporal patterns of the water quality in estuarine ecosystems can reflect the spatial and temporal heterogeneity of these controlling factors. Sediments in shallow waters can impact the physical and ecological environment of a water column through sediment re-suspension and transport (Wang et al. 2008). High TSS concentrations can cause significant mortalities of oysters during the earliest phase of larval development, and further result in decreased oyster filtration rates (e.g. Deksheniaks et al. 2000; Wang et al. 2008). Field data measurements of sediment concentrations are generally expensive or have limited spatial coverage, especially for a large area of estuarine waters (Zawada et al. 2007).

Ritchie et al. (1976) found a curvilinear relationship between suspended sediments and water surface radiance or reflectance (0–200+ mg/L) because the amount of reflected radiance tends to saturate as suspended sediment concentrations increase. Han and Rundquist (1994) investigated the relationships between surface spectral reflectance and suspended sediment concentrations using spectroradiometer and quantum-sensor data in an 8543 L vinyl pool, under natural sunlight, and found the association between surface spectral reflectance and suspended sediment concentrations was linear at low levels (about 600 mg/L) and non-linear at high level of suspended sediment concentrations. The utility of MODIS 250-m data for analyzing complex coastal waters was also examined in the Northern Gulf of Mexico (e.g. Lake Pontchartrain, Mississippi River Delta, and Mississippi Sound), including the mapping of the concentrations of total suspended matter (0–55 mg/L) using simple dark object subtraction (DOS) procedures of atmospheric correction ($R^2 = 0.89$; $n = 52$; $MSE = 4.74$; Miller and McKee 2004). Koponen et al. (2007) archived a retrieval accuracy ($R^2 > 0.9$) of TSS concentration (2.9–20 mg/L) with MEdium Resolution Imaging Spectrometer (MERIS) data (300 m spatial resolution) while the RMSE is 0.74 mg/L (16%) for TSS concentration in the Gulf of Finland (Baltic Sea). Zawada et al. (2007) showed that satellite-based observations are useful for inferring TSS concentrations using a spectra-matching optimization algorithm between the particle backscattering coefficient at 400 nm and TSS (2–50 mg/L) concentrations in Chesapeake Bay from Sea-viewing Wide-Field-of-view Sensor (SeaWiFS) data with low determination coefficient ($R^2 = 0.4$). Fang et al. (2008, 2009) described two methods of deriving surface suspended solid (SS) distributions. One was from 549 nm spectral reflectance (negative correlation due to TSS dominated by particles-carried organic matter) for Hyperion-based TSS concentration mapping, resulting in reasonable spatial distributions of TSS concentration in the Pearl River Estuary (China); the other was by linear fit between in situ reflectance

ratio R680/R527 and TSS concentrations ($R^2 = 0.65$, range: 7.0–241.1 mg/L) for the salinity range of 1.74–22.12.

6.3.2 Chlorophyll-a Monitoring

In recent years, remote sensing has been successfully applied to measure chlorophyll-a concentrations in estuarine and coastal waters (e.g. Gitelson et al. 2005; Werdell et al. 2009; Zimba and Gitelson 2006). Kallio et al. (2001) tested the suitability of the airborne imaging spectroradiometer (AISA) for monitoring lake water quality in four surveys carried out in southern Finland during 1996–1998; they found the ratios L685–691/L670–677 and L685–691/L677–685 gave higher R^2 (0.93 and 0.95, respectively) when the chlorophyll-a concentrations were less than 21 mg/m³. Han and Jordan (2005) developed algorithms for estimating chlorophyll-a concentration (1.14–23.2 mg/m³) in Pensacola Bay using Landsat 7 ETM+ data; their results indicated that the ratio of ETM+ 1/ETM+ 3 was the most effective in estimating chlorophyll-a ($R^2 = 0.67$; standard error (SE) = 1.55 mg/m³). Fang et al. (2008) studied the spatial change tendency in the Pearl River Estuary (China) from the in situ data, and found the negative linear relationship between chlorophyll-a concentration and colored dissolved organic matters (CDOM) absorption under a salinity of 10. The peak value of chlorophyll-a concentration of surface water in tide-way reach was observed where the saltwater and freshwater interface. Originally, a conceptual model was developed and used for estimating pigment concentration of terrestrial vegetation (Gitelson et al. 2003, 2005). Using 14–0.4 ha earthen production ponds having an average depth of 0.9 m, Zimba and Gitelson (2006) established strong linear relationships between analytically measured chl-a (107–3000 mg/m³) and both the three-band $[R-1(650) - R-1(710)] \times R(740)$ and the reflectance ratio model $R(714)/R(650)$; the three-band model accounted for 7% more variation of chl-a concentration than the ratio model (78 vs. 71%). Gitelson et al. (2007) again established strong linear relationships between analytically measured chl-a and both the three-band model $[R-1(675) - R-1(695)] \times R(730)$ and the two-band model $R(720)/R(670)$; the three-band model accounted for 81% of variation in chl-a and allowed estimation of chl-a with a root mean square error (RMSE) of less than 7.9 mg/m³, whereas the two-band model accounted for 79% of chl-a variability and RMSE of estimation was below 8.4 mg/m³. The three-band model was also calibrated and validated using three MERIS spectral bands (660–670, 703.75–713.75, and 750–757.5 nm), and the 2-band model was tested using two MODIS spectral bands ($\lambda_1 = 662–672$, $\lambda_2 = 743–753$ nm) (Gitelson et al. 2008); under turbidity of 1.3–78 NTU, the (1.2–236 mg/m³) predicted by the three-band algorithm was strongly correlated with observed ($R^2 = 0.96$), with a precision of 32% and average bias across data sets of –4.9 to 11%, and predicted by the two-band algorithm was also closely correlated with observed ($R^2 = 0.92$); however, the precision declined to 57%, and average bias across the data sets was 18–50.3%. Le et al. (2009) validated the applicability of a semi-analytical three-band algorithm in estimating concentration in the highly turbid, widely variable waters of Taihu Lake, China,

and improved the algorithm using a proposed four-band algorithm. The improved algorithm is expressed as $[\text{Rrs}(662)-1-\text{Rrs}(693)-1][\text{Rrs}(740)-1-\text{Rrs}(705)-1]-1$. The two semi-analytical algorithms were calibrated and evaluated against two independent datasets collected from 2007 (4–158 mg/m³) and 2005 (0.98–89.2 mg/m³) in Taihu Lake. In several cases, fluorescence line height has been successfully used for the remote detection of in case 2 waters (Gower et al. 1999).

Gons et al. (2008) assessed Chlorophyll-a concentrations and “water-leaving” reflectance along transects in Keweenaw Bay (Lake Superior) and in Green Bay (Lake Michigan) (two of the Laurentian Great Lakes, USA), featuring oligotrophic (0.4–0.8 mg/m³) and eutrophic to hyper-eutrophic waters (11–131 mg/m³); a quite strong linear relationship ($R^2 > 0.81$; standard error (SE) < 0.049 mg/m³) was found between concentration and fluorescence line height (FLH) computed with these MERIS bands, respectively.

6.4 Remote Sensing of Water Quality in Apalachicola Bay

Remote sensing can provide spatial and temporal water quality data covering large areas for estuarine and coastal ecosystem studies. Traditionally, insufficient budget for expensive field data collections has been a major obstacle in ecosystem studies. By using the algorithms originally developed for open ocean waters, NASA provides estimates from SeaWiFS data with 1-km spatial resolution for open ocean. Nevertheless, the 1-km resolution estimates from SeaWiFS are too coarse that are virtually useless for small estuarine ecosystems (such as Apalachicola Bay) with highly optical complexity. On the other hand, MODIS imagery provides much higher spatial resolution, and therefore is suitable to coastal ecosystems that are relatively small. In recent years, we conducted several studies that aimed in monitoring TSS and in Apalachicola Bay by using MODIS data.

6.4.1 Estimating TSS Concentrations

For non-hurricane conditions, Wang et al. (2009) developed a regression model to estimate TSS concentrations based on remote sensor data and two field data sets in Apalachicola Bay, which can be further enhanced by including more data for applications to extreme weather conditions. For investigating the impact of Hurricane Frances on TSS concentrations on Apalachicola Bay, Chen et al. (2009) used an in-water empirical method of atmospheric correction that substantially enhanced the TSS model developed by Wang et al. (2008) for non-hurricane conditions. The remote sensing model used 250-m Moderate Resolution Imaging Spectroradiometer (MODIS) to estimate TSS concentrations in the Bay. The distribution of TSS concentrations in the shallow Gulf of Mexico estuaries was spatially and temporally variable, especially during passage of Hurricane Frances (the average TSS and maximum concentration about 54.3 and 165 mg/L in the Bay, respectively) compared to

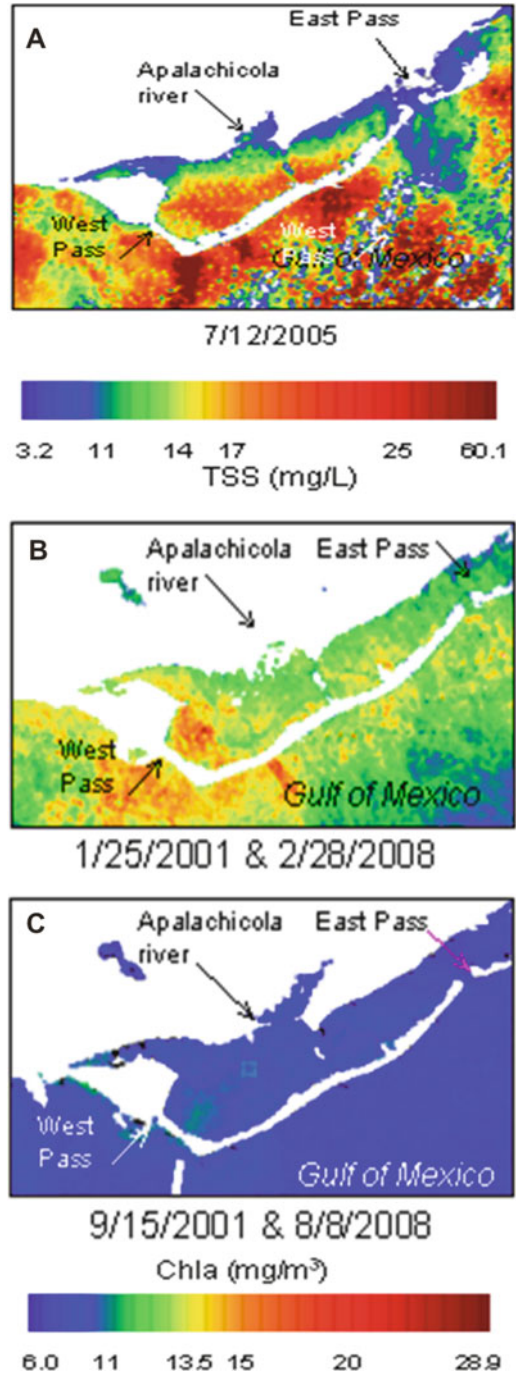
under no-storm or hurricane condition (the average TSS and maximum concentration approximately 24–27 and 58–64 mg/L) (Chen et al. 2009). Chen et al. (2009) also found the southwestward wind (about 50° from north) during the Hurricane Frances induced southwestward currents and transport that resulted in the high TSS concentrations near West Pass in Apalachicola Bay and the Gulf of Mexico. Common to these studies was to relate remote sensor reflectance measured in the red band (600–700 nm) to parameters of water column sediment concentrations. This rationale lies on the dominated scattering from suspended materials when compared to other elements such as colored dissolved organic matter and phytoplankton absorption. The fine spectral characteristics of the MODIS Terra/Aqua instrument plus provide data well suited for the study of TSS in complex coastal waters. The MODIS data that are freely available and at moderately high resolution are useful for estimating TSS concentration in small estuaries.

Using MODIS data and the regression model obtained by Chen et al. (2009), we found relatively high TSS concentrations in Apalachicola Bay as the results of wind-induced sediment resuspension during the Hurricane Frances in 7/21/2005 (Fig. 6.2a). The mean TSS concentration was 17.2 mg/L. The imagery shows the large spatial variations of TSS concentrations. The maximum concentration was 60.1 mg/L, while the minimum concentration was 3.2 mg/L. The standard deviation was 7.0 mg/L.

6.4.2 Estimating Chlorophyll-a Concentrations

Wang et al. (2009) developed a regression model of logarithmically transformed concentration and the ratio of Band 2: Band 1 (MODIS 250 m) to mapping chlorophyll-a concentration from MODIS imagery for Apalachicola Bay ($R^2 = 0.53$, $N = 16$, $P < 0.001$). Huang et al. (2009) applied the regression model developed by Wang et al (2009) to investigate river inflow effects on chlorophyll-a concentrations in the bay for year 2001 and 2008. Averaged high chlorophyll-a concentrations are given in Fig. 6.2b and c for high-flow season (B) and low-flow season (C). Results indicate that chlorophyll-a concentrations were much higher in high-flow season than those in low-flow season. Large spatial differences in Chlorophyll-a concentrations also show an advantage of MODIS-250 m data, which have much higher resolution when comparing with SeaWiFS. Due to wind and tidal effects, higher waters in the bay was exported to low water in the Gulf of Mexico in West Pass, and low ocean waters intruded into the Bay from the East Pass. In general, chlorophyll-a concentration in most areas in the Bay in low-flow season are below 10 mg/L. In high-flow season, the maximum concentration reaches about 18 mg/L. Strong concentrations near inside and outside of West Pass are the evidence of the flux of high-concentration bay water to the Gulf of Mexico through the West Pass inlet. Away from the bay, Chlorophyll-a concentrations were lower in the Gulf because of the dilutions by the less-concentrated waters.

Fig. 6.2 Estimating water quality indicators using MODIS data for Apalachicola Bay, USA. (a) The TSS concentrations derived from the regression model originally developed by Chen et al. (2009). (b and c) Examples of averaged Chl-a concentrations at high flow (b) and low flow (c) conditions



6.5 Summary

Due to large spatial variations, traditional approaches for large-scale spatial and temporal data collections of water quality in estuaries are often very expensive. This paper introduces a cost-effective approach to monitor water quality in estuarine ecosystems by using remote sensing. Among many satellite remote sensing data, MODIS imagery is available at a daily basis with 250 m resolution, which is much higher than the 1-km resolution SeaWiFS products. Although Landsat TM/ETM+ may provide data with much higher spatial resolution, its 27-day temporal resolution makes it difficult to monitor daily variations concerning estuarine dynamics. Remote sensing-based studies of estuarine water quality in Apalachicola Bay have demonstrated the advantages of MODIS data. Our case studies have shown that remote sensing can be used to assess hurricane impacts on estuarine sediment resuspensions and TSS concentrations; remote sensing can also be used to evaluate river inflow effects on chlorophyll-a in estuarine waters. Because chlorophyll-a and TSS are important indicators for water quality and estuarine ecosystem, results from this study can be used to support the environmental and ecological research as well as water resources management in Apalachicola Bay.

References

- Chen SS, Huang WR, Wang HQ, Li D (2009) Remote sensing assessment of sediment resuspension during Hurricane Frances in Apalachicola Bay, USA. *Remote Sens Environ* 12:2670–2681
- Deksheniaks MM, Hofmann EE, Klinck JM, Powell EN (2000) Quantifying the effects of environmental change on an oyster population: a modeling study. *Estuaries* 23:593–610
- Fang LG, Chen SS, Zhang LX, Zeng YH (2008) EO-1 hyperion and ALI inter-sensor comparison of retrieval river suspended solid concentration in Pearl River Estuary, Hydrological Sciences for Managing Water Resources in the Asian Developing World. IAHS-AISH Publication (No. 319):337–346
- Fang LG, Chen SS, Li D, Li HL (2009) Use of reflectance ratios as a proxy for coastal water constituent monitoring in the Pearl River Estuary. *Sensor* 9:656–673
- Gitelson, AA, Gritz U, Merzlyak MN (2003) Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J Plant Physiol* 3:271–282
- Gitelson AA, Viña A, Ciganda V, Rundquist DC, Arkebauer TJ (2005) Remote estimation of canopy chlorophyll content in crops. *Geophys Res Lett* 32, L08403. doi: 10.1029/2005GL022688
- Gitelson AA, Schalles JF, Hladik CM (2007) Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sens Environ* 4:464–472
- Gitelson AA, Dall’Olmo G, Moses W, Rundquist DC, Barrow T, Fisher TR, Gurlin D, Holz J (2008) A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sens Environ* 9:3582–3593
- Gons HJ, Auer MT, Effler SW (2008) MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes. *Remote Sens Environ* 11:4098–4106
- Gower JFR, Doerffer R, Borstad GA (1999) Interpretation of the 685 nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. *Int J Remote Sens* 20:1771–1786

- Han L, Jordan KJ (2005) Estimating and mapping chlorophyll-a concentration in Pensacola Bay, Florida using Landsat ETM+ data. *Int J Remote Sens* 23:5245–5254
- Han L, Rundquist DC (1994) The response of both surface reflectance and underwater light field to various levels of suspended sediments: preliminary results. *Photogramm Eng Remote Sens* 12:1463–1471
- Huang W (2009) Hydrodynamic modeling and ecohydrological analysis of river inflow effects on Apalachicola Bay, Florida, USA. *Estuar Coast Shelf Sci*. doi:10.1016/j.ecss.2009.07.032
- Huang W, Jones K, Wu T (2002) Modeling surface wind effects on subtidal salinity in Apalachicola Bay. *Estuar Coast Shelf Sci* 1:33–46
- Huang W, Chen S, Yang X, Wang H (2009) Remote sensing analysis of river flow effects on chlorophyll-a Concentration in Apalachicola Bay Ecosystem, Florida, USA. Presentation in the 2nd International Conference on Earth Observation for Global Changes, Chengdu, Sichuan, China, May 25–29, 2009
- Kallio K, Kutser T, Hannonen T, Koponen S, Pulliainen J, Vepsäläinen J, Pyhälähti T (2001) Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons. *Sci Total Environ* 1–3:59–77
- Koponen S, Attila J, Pulliainen J, Kallio K, Pyhälähti T, Lindfors A, Rasmus K, Hallikainen M (2007) A case study of airborne and satellite remote sensing of a spring bloom event in the Gulf of Finland (Baltic Sea). *Cont Shelf Res* 2:228–244
- Le CF, Li YM, Zha Y, Sun D, Huang CC, Lu H (2009) A four-band semi-analytical model for estimating chlorophyll a in highly turbid lakes: The case of Taihu Lake, China. *Remote Sens Environ* 6:1175–1182
- Livingston RJ (1984) The ecology of the Apalachicola Bay system: an estuarine profile. U.S. Fish and Wildlife Service, Office of Biological Services. FWS/OBS-82/05
- Livingston RJ, Lewis FG, Woodsum GC, Niu XF, Galperin B, Huang W, Christensen J, Monaco M, Battista T, Klein J, Howell IV, Ray GL (2000) Modeling oyster population response to variation in freshwater input. *Estuar Coast Shelf Sci* 50:655–672
- Miller RL, McKee BA (2004) Using MODIS Terra 250-m imagery to map concentrations of total suspended matter in coastal waters. *Remote Sens Environ* 1–2:259–266
- Ritchie JC, Schiebe FR, McHenry JR (1976) Remote sensing of suspended sediments in surface water. *Photogramm Eng Remote Sens* 2:1539–1545
- Wang HQ, Huang W, Harwell MA, Edmiston L, Johnson E, Hsieh P, Milla K, Christensen J, Stewart J, Liu X (2008). Modeling oyster growth rate by coupling oyster population and hydrodynamic models for Apalachicola Bay, Florida, USA. *Ecol Model* 211:77–89
- Wang HQ, Hladik CM, Milla K, Huang WR, Edmiston L, Harwell MA, Schalles JF (2009). Detecting and mapping water quality indicators in Apalachicola Bay, Florida using MODIS Terra 250-m imagery. *Int J Remote Sens* 31(2):439–453
- Werdell PJ, Bailey SW, Franz BA, Harding LW Jr, Feldman GC, McClain CR (2009) Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua. *Remote Sens Environ* 6:1319–1330
- Whitfield WK Jr, Beaumariage DS (1977). Shellfish management in Apalachicola Bay: past, present, and future. In: Livingston RJ, Joyce EA Jr (eds) Proceedings of the Conference on the Apalachicola Drainage System, Florida Department of Natural Resources Marine Research, Florida
- Zawada DG, Hu CM, Clayton T, Chen ZQ, Brock JC, Muller-Karger FE (2007) Remote sensing of particle backscattering in Chesapeake Bay: A 6-year SeaWiFS retrospective view. *Estuar Coast Shelf Sci* 3–4:792–806
- Zimba PV, Gitelson A (2006) Remote estimation of chlorophyll concentration in hyper-eutrophic aquatic systems: Model tuning and accuracy optimization. *Aquaculture* 1–4:272–286

Chapter 7

Extracting Cryospheric Information over Lowlands from L-Band Polarimetric SAR Data

Nicolas Longép , Masanobu Shimada, Sophie Allain, and Eric Pottier

Abstract Seasonal frozen ground and snow cover are sensitive indicators of how our home planet is changing. In the meantime, new spaceborne SAR systems have been launched, such as the polarimetric PALSAR sensor on-board ALOS in January 2006. In this paper, the relevance of L-band polarimetric SAR data for extracting cryospheric information is presented over lowlands. It is first demonstrated that dry snowpack over frozen ground slightly affects the polarimetric signature. Given the fact that PALSAR data do not enable the use of a simplistic threshold-based method, a refined method for snow detection in PALSAR time series is outlined. A supervised Support Vector Machine is used showing fairly good results within the framework of a three-class classification (dry snow over frozen ground, wet snow and free of snow). Beyond these qualitative studies, a polarimetric EM backscattering model over snow-covered frozen fields brings out the possibility for quantitative assessments. The residual liquid water content in frozen ground over lowlands is estimated from PALSAR measurements.

7.1 Introduction

The cryosphere is an important part of the Earth system, being so interconnected with other parts. Recent changes in the cryosphere have had a major impact on global climate. Seasonal snow cover and frozen ground appear to be some natural barometers of global warming. Over the last two decades, satellite remote sensing has opened the possibility to monitor these cryospheric elements at global scale. Synthetic Aperture Radar (SAR) imaging systems are very promising since they overcome the night-time limitation of optical cameras and the cloud-cover. In addition, microwave SAR sensor can provide valuable information on the vertical structure of the observed natural media. For example, Electromagnetic (EM) waves

N. Long p  (✉)

Earth Observation Research Center, Japan Aerospace Exploration Agency, Sengen 2-1-1,
305-8505 Tsukuba, Ibaraki, Japan
e-mail: nicolas.longepe@jaxa.jp

at L-band (around 1.3 GHz) penetrate dry snow and can provide quantitative information on the underlying media. However, the relationship between the EM backscattering and the ground properties is not straightforward.

In the last 4 years, the Phased Array type L-band Synthetic Aperture Radar sensor (PALSAR) on-board the Advanced Land Observing Satellite (ALOS) has played a key role for the on-orbit demonstration of polarimetric SAR applications (Rosenqvist et al. 2007). Polarimetry can provide interesting information about the intrinsic physical properties of natural media, and was successfully tested for forestry (Hoekman and Quiriones 2000) or crop monitoring (Skriver et al. 2005). Nevertheless, its relevance for extracting cryospheric information over lowlands has not been extensively proved up till now. In this paper, polarimetric PALSAR data are used and their usefulness for seasonal snow and frozen ground monitoring is demonstrated.

Section 7.2 sketches the relation between the polarization of the backscattered waves received by PALSAR sensor and two different cryospheric states (dry snow over frozen ground and wet snow). Neural networks, fuzzy iterative classifier or statistical segmentation have been developed to classify natural media using Polarimetric SAR data. Section 7.3 shows the need to use such a refined method in order to detect these cryospheric states. The Support Vector Machine has been chosen in the framework of this study. Beyond this qualitative assessment, Section 7.4 outlines a method for the retrieval of residual liquid water content in frozen ground. This method is based on the simulation of a snow EM model adapted to this L-band case.

7.2 Polarimetric Studies over Snow-Covered Agricultural Fields

7.2.1 Data

PALSAR is a fully polarimetric instrument, operating at L-band (1270 MHz center frequency) with a 21.5   default off-nadir angle for polarimetric acquisitions. This mode provides the full polarimetric scattering matrix (S_{HH} , S_{HV} , S_{VH} , S_{VV}) over a 30 km width swath with a 9.36 m \times 4.45 m slant range resolution. Some calibration characteristics of the PALSAR instrument are given in Table 7.1, with the corresponding pre-launch specifications. These results summarize the initial calibration phase, which covers five months between May 16, 2006 and October 23, 2006 and the first half-year of the operational phase. PALSAR data were calibrated and validated using a total of 500 calibration points collected worldwide and distributed target data over the Amazon (Shimada et al. 2009). Since the PALSAR radiometric performances are better than the pre-launch specifications, PALSAR sensor is stable enough to perform temporal polarimetric analysis.

Even though the effects of terrain relief on polarimetric SAR data can be significantly reduced (Schuler et al. 1999), the variability of incidence angles

- Wet snow on March 2007: Snow depth, temperature, rainfall and snowfall measurements were performed by automatic stations during 2007 and 2008 winters and collected by the Japanese Meteorological Agency (JMA). PALSAR acquisition on March 30, 2007, occurred during a snow melting period. Prolonged mild spell at the end of March reduced the snowpack from approximately 45 to 15 cm in about 10 days. The snowpack is thoroughly wet in most locations with a good chance of bare spots.
- Bare soil on November 2007: By using polarimetric incoherent decomposition theorems (Cloude and Pottier 1996; Freeman and Durden 1998), data acquired on November 15, 2007, show the lowest volume scattering contribution. At this date, the ground is still unfrozen and no snowfall was measured by the JMA: the fields can be considered as “bare soil”. It is in agreement with the aforementioned sowing and harvest periods.
- Dry snow over frozen ground on February 2008: During the last polarimetric PALSAR/ALOS overpass on February 15, 2008, ground-truth measurements were conducted within the test site indicating completely dry low-density snowpack over frozen soil. At L-band, this snowpack can be considered as spatially homogeneous with a 30 cm thickness and a density of 0.21 ± 0.01 . Hirota et al. (2006) investigated the relationship between daily temperature and frost depth over this area, and proposed the following equation:

$$D_{\max} \approx \beta \sqrt{F_{20}} \quad (7.1)$$

where D_{\max} is the maximum frost depth and β an empirical coefficient. The soil freezing index F_{20} sums days with a daily average air temperature below 0°C until snow depth reaches 20 cm. For this acquisition date in February, frost depth is estimated to be equal to 14 cm.

Finally, it is worth emphasizing that the five other acquisitions are realized in the presence of agricultural vegetation and will be used to point out the uniqueness of the three aforementioned ones.

7.2.2 Qualitative Polarimetric Analysis

In the case of natural media, polarimetric data are generally processed using incoherent averaging techniques and speckle reduction. In order to characterize the second order statistical properties of a set of n independent matrices, n -look 3×3 coherency matrix \mathbf{T} is built as follows:

$$\mathbf{T} = \frac{1}{n} \sum_{i=1}^n \mathbf{k}_i \mathbf{k}_i^{*T} \quad \text{with } \mathbf{k} = \frac{1}{\sqrt{2}} [S_{\text{HH}} + S_{\text{VV}}, S_{\text{HH}} - S_{\text{VV}}, 2S_{\text{HV}}]^T \quad (7.2)$$

with \mathbf{k} the so-called target scattering vector in the Pauli polarization basis depending on the scattering complex element S_{pq} . In this study, a 4-look averaging in azimuth

with a Lee refined speckle filtering (Lee et al. 1999) are applied on these polarimetric PALSAR data. Numerous descriptors extracted from the coherency measured matrix can be analyzed in order to find the best thresholds that might split data into different classes. In Fig. 7.2, some histograms are plotted for various relevant polarimetric descriptors acquired over the agricultural fields. The EM backscattering coefficients (HH, VV and VH channels), the Freeman–Durden powers (the single bounce P_s and volume P_{vol} components) and the Cloude–Pottier eigenvalue-eigenvector parameters (the entropy H , $\bar{\alpha}$ angle and anisotropy A)¹ are illustrated. The aim of this qualitative study is to analyze the influence of snow cover and frost on the polarimetric signature. Consequently, these histograms have been normalized against data acquired over bare soil condition (November 2007). The understanding of scattering mechanisms is a delicate issue and largely depends on the chosen polarimetric parameters. However, the following trends can be observed:

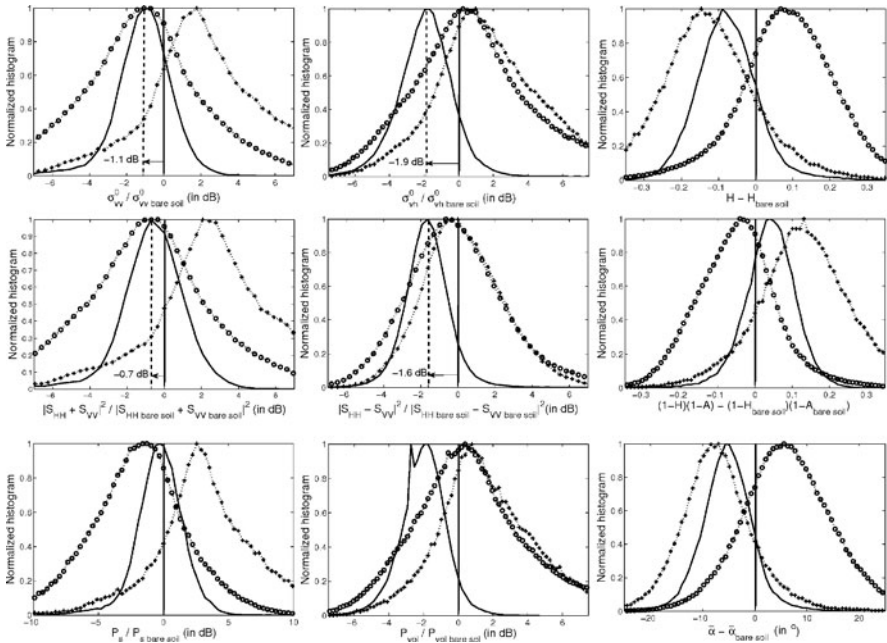


Fig. 7.2 Histograms of various polarimetric descriptors over agricultural fields with different cryospheric conditions compared to the bare soil case: wet snow on March 2007 (+), dry snow over frozen ground on February 2008 (—) and soil with small vegetation using the five other acquisitions (o). The significance of the shift values will be discussed in Section 7.4.2

¹The entropy determines the degree of randomness of the scattering process, the $\bar{\alpha}$ angle is related to the physics behind the scattering process ($\bar{\alpha} \rightarrow 0$: single-bounce, $\bar{\alpha} \rightarrow \pi/4$: volume, $\bar{\alpha} \rightarrow \pi/2$: double-bounce). The anisotropy A measures the relative importance of the second and the third eigenvalues of the eigen decomposition.

- Case of wet snow: The co-polarized backscattering coefficients σ_{VV}^0 and σ_{HH}^0 (not shown) are higher in the case of wet snow cover, and the Freeman–Durden decomposition shows an increase in the single bounce phenomena (left column). In agreement with these observations, the diminutions of the entropy H and the $\bar{\alpha}$ angle confirm the increase of the single bounce contribution with wet snow cover.
- Case of dry snow with frozen ground: A slight difference can be observed between the co-polarized backscattering coefficients of dry snow with frozen ground and the ones of bare soil. Histograms are Gaussian-shaped centered around -1.10 dB (σ_{VV}^0) and -0.54 dB (σ_{HH}^0 -not shown) with approximately 1.6 dB standard deviation. By contrast, some polarimetric parameters provide useful information (central column). Dry snow cover over frozen ground induces a decrease of secondary mechanisms, as suggested by P_{vol} or σ_{VH}^0 .
- Case of vegetation cover: As expected, some polarimetric indicators such as the entropy H or the $\bar{\alpha}$ angle clearly indicate that the volume contribution is higher in case of crops. This point clearly shows that polarimetric SAR data are suitable for crop monitoring.

Polarimetric SAR data acquired under dry or wet snow condition at L-band possess specific attributes and seem to be identifiable in time series. However, it should be noted that the standard deviation for all of these histograms does not allow the use of simplistic threshold-based method. A refined methodology should be outlined.

7.3 State of the Cryosphere by Means of Statistical Learning Method

In the previous section, it was qualitatively shown that the state of the cryosphere over lowlands and polarimetric PALSAR data establish a confusing relation. In another hand, classification algorithms based on statistical learning methods such as the Support Vector Machine (SVM) are used in a wide range of data mining applications. SVM has been successfully introduced in remote sensing (Foody and Mathur 2004; Melgani and Bruzzone 2004) and is investigated in the framework of this study.

7.3.1 SVM Background

In this paper, SVM theory is outlined for binary classification. Further explanations can be found in numerous books, especially Vapnik (1998) and Cristianini and Shawe-Taylor (2000). First, training SAR data $x \in \mathbb{R}^n$ are assigned to one of the two possible clusters $\Theta_i \in \{+1; -1\}$. For each pixel, p stands for the number of elements derived from the polarimetric measurements (backscattering coefficients, channel correlation, parameters derived from incoherent decompositions. . .). In this Euclidean space of dimension p , SVM computes a decision function (hyperplane) which separates the positive from the negative samples. The decision function can

be written as $f(\mathbf{x}) = \mathbf{w}\mathbf{x} + b$ where $\mathbf{x}_i \in \mathbb{R}^p$ is the i^{th} sample vector, $\mathbf{w} \in [\mathbb{R}]^p$ a weight vector and b a scalar. An equivalent problem consists in solving the following constraints:

$$\left. \begin{aligned} \mathbf{x}_i\mathbf{w} + b &\geq +1(-\xi_i^*) \text{ for } \Theta_i = +1 \\ \mathbf{x}_i\mathbf{w} + b &\leq -1(+\xi_i^*) \text{ for } \Theta_i = -1 \end{aligned} \right\} \quad (7.3)$$

After some manipulations, the distance between the closest positive and negative samples is equal to $2/\|\mathbf{w}\|$ where $\|\mathbf{w}\|$ is the Euclidean norm of \mathbf{w} . Consequently, the optimal hyperplane is given by maximizing this margin, subject to the constraints defined by Eq. (7.3). The solution is illustrated on the top panel of Fig. 7.3.

In general, input data are non-separable and positive slack variables ξ should be introduced in the aforementioned constraints – see * in Eq. (7.3). ξ_i is a measure of the misclassification error for the i^{th} sample vector and exceeds unity in that case. Therefore, the generalized optimal separating hyperplane is determined by the vector \mathbf{w} subject to the constraints of Eq. (7.3), minimizing:

$$\Phi(\mathbf{w}, \xi) = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_i \xi_i \quad (7.4)$$

where C is a user-defined parameter accounting for the relative importance of the misclassification errors. In the case where linear boundary is inappropriate, data

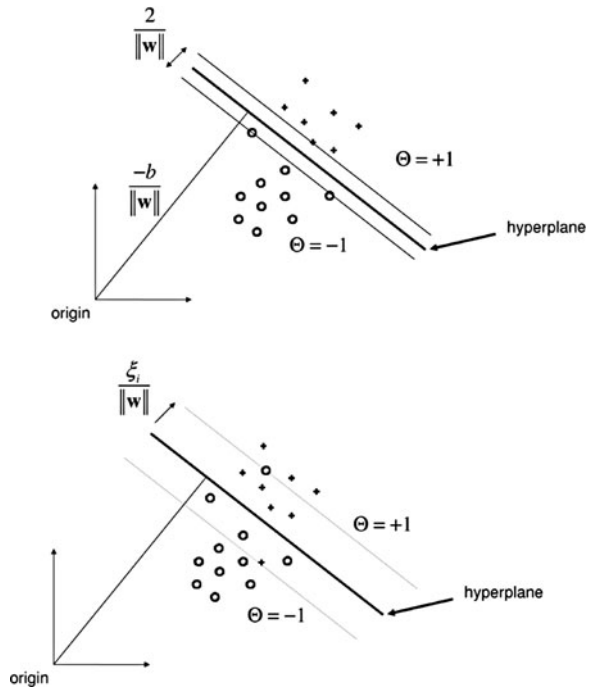


Fig. 7.3 Linear separating hyperplanes for the separable case (top) and non-separable case (bottom) by Burges (1998)

are projected through a ‘‘Kernel function’’ to another Euclidean space with higher dimension where SVM can find a linear separating hyperplane. Choosing non-linear kernels allows to construct classifiers that are linear in the feature space, even though they are non-linear in the original space. This projection is very powerful compared to the classical threshold-based classification. As a default kernel, the Gaussian Radial Basis Function (RBF) of the form

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2} \quad (7.5)$$

is usually used. It should be noted that the user-defined penalty parameter, C , and the kernel parameters, γ , have to be optimized for each vector configuration by using a ‘‘grid-search’’ with k -fold cross-validation.²

7.3.2 Methodology

In the framework of this study, a three-class classification (‘‘dry snow over frozen ground’’, ‘‘wet snow’’ and ‘‘free of snow’’) is realized based on the eight polarimetric PALSAR acquisitions and the SVM method. Following Section 7.2.1, data acquired on February 2008 (March 2007, resp.) are considered as belonging to the ‘‘dry snow over frozen ground’’ class (‘‘wet snow’’ class, resp.). The six others among which the one acquired under bare soil condition are labeled as ‘‘free of snow’’. The same number of training data is randomly selected for each class (i.e. 2500 pixels). For the ‘‘snow free’’ class, training data are equiprobably chosen among the 6 acquisition dates. Numerical implementation by Chang and Lin (2001) is used with the RBF kernel. The classifier is optimized finding the best (C, γ) pair using a fivefold cross-validation.

The capabilities of polarimetric PALSAR data for detecting the state of the cryosphere are then tested over more than 250,000 pixels for each acquisition. The assessment of classification accuracy is based on the confusion matrix, which is derived after comparison of estimated classes against true classes. Four different SVM tests are presented in this study depending on the input vector (the three backscattering coefficient, the Freeman–Durden decomposition powers, the Cloude–Pottier eigenvector-based parameters or the coherency matrix). Finally, the classical classification method based on the Maximum Likelihood (ML) with a multivariate complex Wishart distribution of coherency matrices (Lee et al. 1994) is tested in this study. The results can be directly compared to the ones issued from the SVM with the nine elements of the coherency matrices.

² k -fold cross-validation consists in partitioning the data set into k subsets. Of the k subsets, $k - 1$ subsets are used as training data, and the remaining single subset is retained as the validation data for testing the hyperplane. The cross-validation process is then repeated k times, with each of the k subsets used exactly once as validation data. Finally, the k results are averaged producing a single and stable estimation.

Table 7.2 Classification accuracies (in %) using SVM with different polarimetric parameters and classical ML decision with a multivariate complex Wishart distribution of coherency matrices

	Estimation			Overall
	Free	Wet	Dry	Accuracy
SVM with Backscattering powers				
Free of snow	68.9	5.8	25.4	68
Wet snow	11.6	50.1	38.3	
Dry snow/frozen ground	15.4	5.7	78.9	
User accuracy	94	55	29	
SVM with Freeman powers				
Free of snow	81.5	6.9	11.6	83
Wet snow	10.6	76.6	12.8	
Dry snow/frozen ground	4.1	1.2	94.7	
User accuracy	97	64	53	
SVM with H/A/α powers				
Free of snow	75.1	6.4	18.5	73
Wet snow	10.5	53.5	36	
Dry snow/frozen ground	8.1	15.3	76.7	
User accuracy	96	50	34	
SVM with Coherency matrix \mathbf{T}				
Free of snow	76.2	6.2	17.6	74
Wet snow	13.3	52.8	34	
Dry snow/frozen ground	12.6	6.2	81.1	
User accuracy	95	55	37	
ML with Wishart \mathbf{T}				
Free of snow	85.0	5.6	9.4	74
Wet snow	41.7	41.1	17.3	
Dry snow/frozen ground	47.5	13.7	33.8	
User accuracy	85	46	35	

The confusion matrices are showed in Table 7.2. The diagonal elements represent correctly classified pixels while the cross-diagonal elements represent misclassified pixels. The User Accuracy indicates overestimation and is defined for each class as the ratio between the number of correct classifications and the total number of classifications in the category.

7.3.3 Results

First of all, we can notice that the overall classification accuracy of SVM with the three backscattering coefficients is significantly lower than others. As shown in the last section, backscattering coefficients are not specifically linked

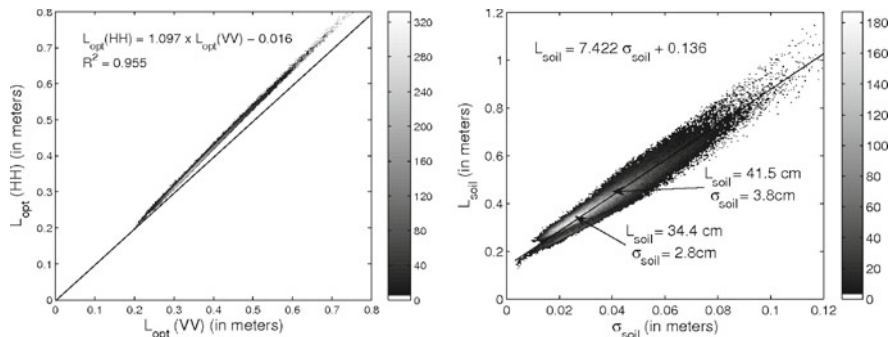


Fig. 7.4 Correlation between the soil correlation lengths determined with VV and HH channels (*left*). Soil roughness characteristics estimated by L-band PALSAR measurements over the agricultural bare soil on November 2007 (*right*). Gray tones indicate the number of occurrences

to physical mechanisms and do not exclusively depend on cryospheric information. SVM with Freeman powers shows the best accuracy (OA = 83%) revealing the relevance of this polarimetric decomposition in providing complementary powers. Other input parameters give fairly good results with overall accuracies around 74%.

It should be noted that wet snow detection seems difficult as compared to the “dry snow and frozen ground” case. This result is unexpected since wet snow should have a stronger impact on the backscattering mechanism. By contrast, the detection of dry snow cover over frozen ground is excellent with Freeman decomposition (94.7%). Depending on the input parameters, this score varies and the false alarm rate for the “dry snow/frozen ground” case is significant resulting in a User Accuracy below 40%. Finally, it is shown that snow detection seems limited with the ML decision: the pseudo-distance based on the Wishart distribution is inappropriate.

In a general manner, polarimetric L-band SAR data provide useful information related to the state of the cryosphere. SVM classifier with appropriate polarimetric parameters derived from PALSAR data seems particularly powerful. The results in the case of SVM with Freeman decomposition are illustrated by Fig. 7.4. However, this temporal representation reveals the fact that the Freeman decomposition in September 2006 is statistically close to the “dry snow/frozen ground” case for the SVM technique. The proposed methodology has to be improved, further optimizing the Kernel function, and finding the most relevant polarimetric parameters using a recursive feature elimination approach.

7.4 Quantitative Assessments

Beyond this qualitative assessment, this section introduces a methodology for extracting quantitative cryospheric parameters over lowlands. Given the uncertainties related to the soil condition in March 2007, this section will focus on the

cryospheric conditions on February 2008. Given the penetration depth of dry snowpack at L-band (over 50 m), the total backscattering power in February is mainly due to the soil contribution through the snowpack. Soil roughness is assumed constant with the absence of plowing between the acquisitions on November 30, 2007, and February 15, 2008. The only significant change about the ground concerns liquid water content. In order to model the impact of dry snow and frost on the polarimetric signature of February, the soil characteristics are first estimated using PALSAR data acquired in November.

7.4.1 Bare Soil Characterization

Rough surface is generally assumed to be defined as a stationary process. It is described by a height probability distribution function (exponential form defined by the Root Mean Square (RMS) surface height σ_{soil}) and a surface correlation function (defined by the correlation length L_{soil}).

The RMS surface height σ_{soil} and the soil moisture M_v can be estimated by the polarimetric Oh's retrieval method (Oh et al. 2004). This method is based on a semi-empirical polarimetric backscattering model developed both with theoretical models (Integral Equation Model (IEM) (Fung et al. 1992) and Geometrical Optics) and an extensive database. From radar observations of σ_{VV}^0 , σ_{HH}^0 and σ_{VH}^0 , the direct model is inverted through an inversion diagram of five different equations resulting in the estimation of these two parameters. This algorithm is applied over the test area with PALSAR data acquired in November 2007. The estimated median value for the soil moisture is equal to 11.6%.³ With σ_{soil} and M_v estimated, the soil correlation length is then retrieved by the optimization of the IEM simulation in the co-polarized channels regarding PALSAR measurements. Cross-polarized channels are not used due to the uncertain accuracy of the IEM model to simulate multiple scattering phenomena. Two different correlation lengths can be estimated according to the chosen channel (VV or HH). The correlation between the retrieved correlation lengths, $L_{\text{opt}}^{\text{VV}}$ and $L_{\text{opt}}^{\text{HH}}$, is plotted on the left panel of Fig. 7.5.

It can be noticed that the correlation is excellent. However, the higher the correlation length is, the more the values retrieved with HH channel are superior to the ones with VV channel. IEM model linearly diverges from co-polarized PALSAR measurements for low backscattering (i.e. high correlation length). This bias remains nonetheless negligible confirming the relevance of the IEM model for simulating the EM backscattering over bare soil. The final correlation length (average of the two optimal L_{opt}) and the RMS height are plotted on the right panel of Fig. 7.5. The regression equation ($L_{\text{soil}} \approx 7.41 \sigma_{\text{soil}} + 0.14$) seems realistic for agricultural bare soil. Soil characteristics retrieved from PALSAR measurements in November 2007 will be used in the following subsections.

³ This value has been confirmed by a hydrologic model proposed by Dingman (2002) coupled with meteorological data provided by the JMA and soil information provided by Webb et al. (1991).

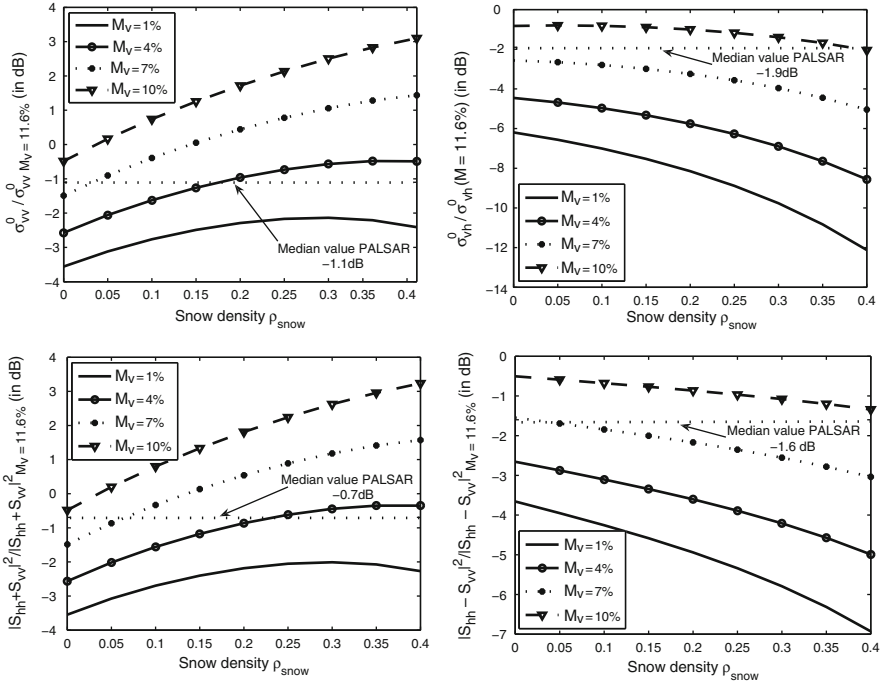


Fig. 7.5 EM backscattering simulation over dry snowpack and frozen ground depending on snow density for different residual liquid water content $M_v \in [1-10]\%$. The values are normalized with the simulated EM backscattering over the ground condition in November 2007

7.4.2 Electromagnetic Backscattering Modeling for Snow-Covered Frozen Ground

The polarimetric signature induced by dry snow cover over frozen ground is simulated in this subsection. Dry snow cover of density ρ_{snow} induces four effects on the EM backscattering at L-band (Shi and Dozier 2000).

- The incident wavelength within the snowpack λ_{snow} is shorter because snow is dielectrically thicker than air ($\epsilon_{snow} > \epsilon_{air}$). The surface appears rougher, inducing an increase of reflectivity.
- The snow near the ground reduces the dielectric contrast $\epsilon_{soil}/\epsilon_{air} \rightarrow \epsilon_{soil}/\epsilon_{snow}$, reducing in turn the backscattering.
- The incidence angle at the snow-ground interface is smaller due to the refraction within the snow.
- Volume contribution and extinction are neglected. No vertical snowpack structure can influence the polarimetric signature.

Finally, the total polarimetric backscattering over a dry snowpack is equal to:

$$\mathbf{T}_{\text{dryshow}} = \frac{\cos \theta_0}{\cos \theta_{\text{snow}}} \mathbf{R}(\theta_{\text{snow}}, \varepsilon_{\text{soil}}/\varepsilon_{\text{snow}}, \lambda_{\text{snow}}, L_{\text{soil}}, \sigma_{\text{soil}}) \quad (7.6)$$

where θ_0 and θ_{snow} are the incidence angles in the air and within the snowpack, respectively. \mathbf{R} represents the 3×3 coherency matrix computed with the IEM proposed by Fung et al. (1992). The hypothesis of reflection symmetry, which is generally valid for rough surface, is used in this study. Correlation between co-polarized and cross-polarized channels is assumed to be null.

$$\langle S_{pp} S_{pq}^* \rangle = \langle S_{qq} S_{pq}^* \rangle = 0 \quad (7.7)$$

Following Eq. (7.6), the EM backscattering over the dry snowpack on February 2008 is simulated having the median soil characteristics estimated with PALSAR data acquired in November 2007 ($L_{\text{soil}} = 34.4 \text{ cm}$, $\sigma_{\text{soil}} = 2.8 \text{ cm}$). Some polarimetric parameters are plotted on Fig. 7.6 as function of snow density and soil moisture content (low value in case of frost). In order to study the impact of these two cryospheric elements, each polarimetric parameter is again normalized by the

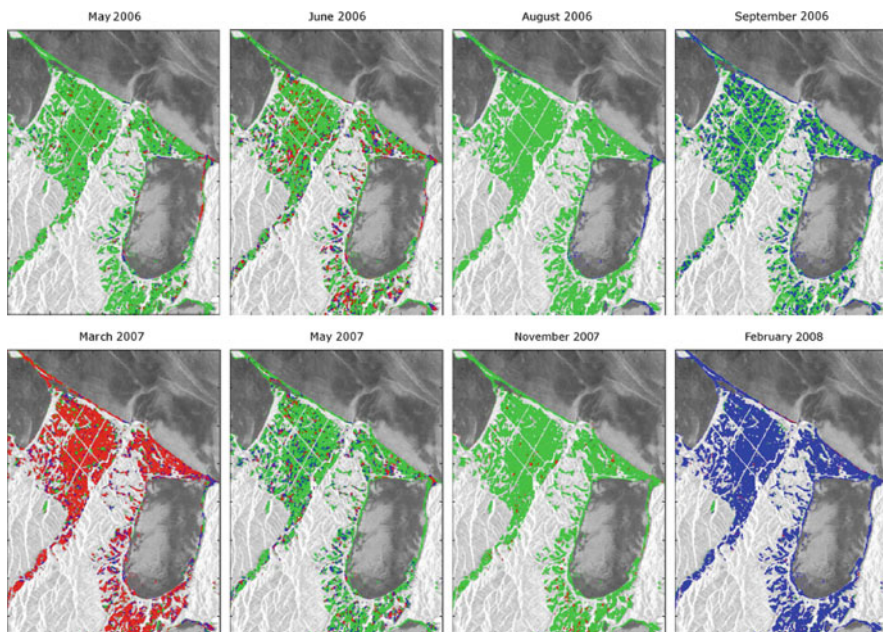


Fig. 7.6 Classification of the cryospheric state over agricultural fields by means of SVM and Freeman decomposition powers derived from PALSAR measurements: *blue* “Dry snow/frozen ground”, *red* “Wet snow” and *green* “Free of snow”. *Gray* tones represent other non-used natural media

simulated parameter over the median bare soil condition found in November 2007 (i.e. with M_v equal to 11.6%).

Lower the soil moisture (e.g. permittivity), lower the EM backscattering intensities, both for the co- and cross-polarized channels. EM backscattering over dry snowpack mainly depends on this parameter at L-band. Snowpack induces also a modification of the polarimetric signature. Single scattering $|S_{HH} + S_{VV}|^2$ increases whereas volume σ_{VH}^0 and double bounce $|S_{HH} + S_{VV}|^2$ phenomena decrease according to the snow density. This point is in agreement with Fig. 7.2. Dry snowpack induces a global decrease of the secondary mechanisms. In a qualitative manner, the modifications due to dry snow and frozen ground are correctly considered with this EM modeling.

Knowing the snow density equal to 0.21 in February 2008, polarimetric changes measured by PALSAR can be fully explained by a soil moisture content ranging from 4% to 9%. The uncertainties on the IEM model for predicting multiple scattering effects and correlation terms $\langle S_{pp}S_{qq}^* \rangle$ may explain this interval. However, these residual liquid water contents are fully realistic for frozen ground.

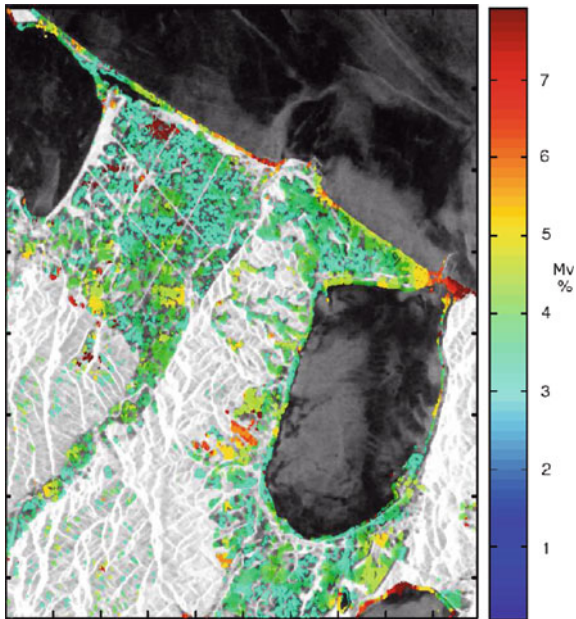
7.4.3 Toward the Estimation of Residual Liquid Water Content in Frozen Ground

Polarimetric signature at L-band over dry snow-covered frozen fields depends on snow density, but furthermore on soil moisture. Given the fact that in situ snow measurements reveal a homogeneous snow cover with a constant density over these flat agricultural fields, the aim of this subsection is to show the feasibility of retrieving the residual liquid water content in frozen ground. For this preliminary study, we assume that the snow density is known and equal to 0.21.

The proposed methodology is based on the aforementioned EM modeling and PALSAR data acquired on February 2008. Given the uncertainties concerning the simulation of cross-polarized channels and co-polarized correlation terms, only the co-polarized backscattering coefficient σ_{VV}^0 is used for this preliminary quantitative assessment. For each pixel, we propose to optimize the soil moisture by matching the simulated and measured PALSAR histograms on February derived from a centered $N_b \times N_b$ pixels area. Over these snow-covered frozen fields, the EM simulations depend also on the snow density assumed homogeneous and on the soil roughness map previously estimated. The Chi-square test is then used to measure the goodness-of-fit of the simulated distributions for the VV channel. Given $\mathcal{V}_{palsar}^{VV}$ and $\mathcal{U}_{sim(M_v)}^{VV}$ two histograms corresponding, respectively, to the PALSAR distribution and the simulated σ_{VV}^0 distribution for a given soil moisture M_v , χ_v^2 statistic is equal to:

$$\chi_v^2 \left(\mathcal{U}_{sim(M_v)}^{VV}, \mathcal{V}_{palsar}^{VV} \right) = \sum_k \frac{(u_k - v_k)^2}{v_k} \quad (7.8)$$

Fig. 7.7 Estimation of the residual liquid water content in frozen agricultural fields on February 15, 2008, by using PALSAR co-polarized VV channel. *Gray* tones represent other non-used natural media



where u_k and v_k are the simulated and measured frequency for bin k , respectively. For each pixel over a centered $N_b \times N_b$ box, the optimal residual soil moisture is found by minimizing the metric:

$$M_{\text{vopt}} = \arg \min_{M_v} \chi_v^2 \left(\mathcal{U}_{\text{sim}(M_v)}^{VV}, \mathcal{V}_{\text{palsar}}^{VV} \right) \quad (7.9)$$

Histogram matching prevents from estimating spatially fluctuating M_{vopt} which might be induced by residual speckle effects or soil roughness modifications at pixel scale. Results are illustrated on Fig. 7.7 using $N_b = 11$.

Even if the accuracy of such a method can not be assessed, the retrieved values seem perfectly realistic with $M_{\text{vopt}} \approx 3 - 5\%$. The main limitation of this method is that the IEM validity is not perfectly guaranteed over frozen ground. The use of an impedance matching approach accounting for surface effects may be of interest in the framework of this quantitative assessment (Schwank et al. 2004). Volume contribution from the soil might also influence the results.

7.5 Conclusions

This paper shows the contribution of polarimetric L-band PALSAR data for extracting cryospheric information over lowlands. The launch of this sensor on-board ALOS satellite represents a real opportunity in the uncertain context of global warming. It appears that the polarimetric signature over snow-covered flat bare

fields presents some interesting features at L-band. SVM classifier seems to be a valuable tool for extracting the uniqueness of the polarimetric measurements over “dry snow/frozen ground” or “wet snow” condition. Finally, a polarimetric L-band snow model based on the Integral Equation Model is implemented and tested for the particular case of frozen ground. Simulations are in relatively good agreement with PALSAR measurements, opening the way for quantitative assessments. A method based on local statistics (χ^2 test) is finally performed in order to retrieve the residual liquid water content in frozen ground. Results seem very promising, even if the polarimetric EM modeling over frozen ground still remains a challenging issue.

This work has been carried out using data acquired by the PALSAR sensor on-board ALOS which have the long revisit time of 46 days. In the design of the PALSAR strategy (Rosenqvist et al. 2007), the full polarimetric mode has been chosen as an experimental mode to promote research and development. Out of the 16 cycles realized in 2007 and 2008, only two of them were fully dedicated to the polarimetric mode. Regarding the methods proposed in this study, a comprehensive knowledge of the natural media and its polarimetric backscattering is necessary to extrapolate these locally developed methods to a regional or global scale context. The forthcoming ALOS-2 satellite with an L-band SAR system on-board (scheduled in 2013/2014) opens new prospects for a better understanding. Its shorter repeat-pass orbit (14 days) with a true spatial and temporal consistency for the polarimetric mode might be the key component for monitoring snow and land ice at global scale. In addition to the proposed methods, Interferometric L-band SAR technique could be a real asset for estimating the distribution of the Snow Water Equivalent.

Acknowledgments The authors greatly acknowledge the Japan Society for the Promotion of Sciences (JSPS) for providing financial support from October 2007 to September 2008 through its short-term fellowship program. The authors would like to thank A. Hachikubo, T. Tanikawa (Kitami Institute of Technology, Japan) and Y. Miyagi (EORC/JAXA) for their help and support during the snow campaign held on February 2008.

References

- Burges CJC (1998) A tutorial on support vector machines for pattern recognition. *Data Min Knowl Disc* 2(2): 121–167
- Chang C, Lin C (2001) Libsvm: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Cloude SR, Pottier E (1996) A review of target decomposition theorems in radar polarimetry. *IEEE T Geosci Remote* 34:498–518
- Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines (and other kernel-based learning methods). Cambridge University Press, Cambridge
- Dingman S (2002) *Physical hydrology*. Prentice Hall, NJ
- Foody GM, Mathur A (2004) A relative evaluation of multiclass image classification by support vector machines. *IEEE T Geosci Remote* 42:1335–1343
- Freeman A, Durden S (1998) A three-component scattering model for polarimetric SAR data. *IEEE T Geosci Remote* 36:963–973
- Fung A, Li Z, Chen K (1992) Backscattering from a randomly rough dielectric surface. *IEEE T Geosci Remote* 30:356–369

- Hirota T, Iwata Y, Hayashi M, Suzuki S, Hamasaki T, Sameshima R, Takayabu I (2006) Decreasing soil-frost depth and its relation to climate change in Tokachi, Hokkaido, Japan. *J Meteorol Soc Jpn* 84(4):821–833
- Hoekman DH, Quiriones MJ (2000) Land cover type and biomass classification using AIRSAR data for evaluation of monitoring scenarios in the Colombian Amazon. *IEEE T Geosci Remote* 38:685–696
- Lee JS, Grunes M, deGrandi G (1999) Polarimetric SAR speckle filtering and its implication for classification. *IEEE Trans Geosci Remote* 37:2363–2373
- Lee J, Grunes M, Kwok R (1994) Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution. *Int J Remote Sens* 15(11):2299–2311
- Melgani F, Bruzzone L (2004) Classification of hyperspectral remote sensing images with support vector machines. *IEEE T Geosci Remote* 42:1778–1790
- Oh Y, Sarabandi K, Ulaby F (2004) Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces. *IEEE Trans Geosci Remote* 42(3):596–601
- Rosenqvist A, Shimada M, Ito N, Watanabe M (2007) ALOS PALSAR: A pathfinder mission for global-scale monitoring of the environment. *IEEE Trans Geosci Remote* 45:3307–3316
- Schuler D, Lee J, Ainsworth T (1999) Compensation of terrain azimuthal slope effects in geophysical parameter studies using polarimetric SAR data. *Remote Sens Environ* 69:139–155
- Schwank M, Stahli M, Wydler H, Leuenberger J, Matzler C, Fluhler H (2004) Microwave L-band emission of freezing soil. *IEEE Trans Geosci Remote* 42:1252–1261
- Shi J, Dozier J (2000) Estimation of snow water equivalence using SIR-C/X-SAR. (I) Inferring snow density and subsurface properties. *IEEE Trans Geosci Remote* 38:2465–2474
- Shimada M, Isoguchi O, Tadono T, Isono K (2009) PALSAR radiometric and geometric calibration. *IEEE Trans Geosci Remote* 47:3915–3932
- Skriver H, Dall J, Toan TL, Quegan S, Ferro-Famil L, Pottier E, Lumsdon P, Moshammer R (2005) Agriculture classification using POLSAR data. In: *Proceedings of the POLINSAR'05, Frascati, Italy*
- Vapnik V (1998) *Statistical learning theory*. Wiley-Interscience, New York
- Webb R, Rosenzweig C, Levine E (1991) A global data set of soil particle size properties. *Tech Rep, NASA Technical memorandum 4286*

Chapter 8

Variability of Northern Hemisphere Spring Snowmelt Dates Using the AVHRR Polar Pathfinder Snow Cover During 1982–2004

Hongxu Zhao and Richard Fernandes

Abstract A new daily 5 km resolution Northern Hemisphere (NH) snow cover product derived from AVHRR Polar Pathfinder data (APP), available from 1982 to 2004, has enabled the investigation of the variability of the spring snowmelt dates at continental and hemispheric scales in association with variability of circumpolar climate parameters. Continentally averaged and latitude band averaged snowmelt date time series were analyzed together with melting season temperatures, preceding winter precipitation and atmospheric circulations. Snowmelt dates over northern Eurasia and North America did not show a statistically significant trend but rather a strong interannual variability related to anomalously large-scale atmospheric circulations. The average North American snowmelt date was observed to be later than that of northern Eurasia except the year of 1998. Snowmelt dates co-varied on interannual timescales between North America and the northern Eurasia before 1998. After 1998, snowmelt dates between the two continents are poorly related. The change in regime may be associated with the interdecadal coupling in climate variability between North Pacific and North Atlantic. Finally, the less contribution in interannual variability of snow factor to snow albedo feedback was explained by analyzing seasonal snow-temperature sensitivity regions.

8.1 Introduction

Snow cover variability is believed to be an important indicator of climate change that also strongly influences regional and global climate (Brown et al. 2007; Groisman et al., 1994a; Qu and Hall, 2007), affects land cover characteristics and land-atmosphere energy exchange, and modifies habitat suitability (Massom 1991). The large-scale spatial distribution and interannual variability of snow cover over NH

H. Zhao (✉)

Adaptation and Impacts Research Section, Climate Research Division, Environment Canada, Toronto, Canada

e-mail: hongxu.zhao@ec.gc.ca

high-latitude lands has attracted considerable attention in recent years because it is acknowledged that the snow albedo feedback is a leading factor of amplified warming in northern high latitude regions (Hall et al. 2008). Many studies indicate that the maximum snow-albedo feedback occurs during spring (Déry and Brown 2007; Groisman et al. 1994a; Hall and Qu 2006). Hence, the possibility of persistent changes in snow cover under recent global warming will play a significant role in the surface radiation budget (Déry and Brown 2007; Qu and Hall 2005). Based on weekly satellite snow cover extent (SCE) over NH continents, Robinson (2003) identified an abrupt decrease in snow cover extent occurred in the spring and early summer since the 1980s. Brown (2000) and Ye et al. (1998) suggested a consistent circumpolar increase in observed snowfall during the winter time. Most recently, based on model simulations with greenhouse gas induced climate change, Raisanen (2008) suggested that variations in snow amount at a geographical point depend on both warming during winter and spring and increased snowfall. Hence, changes in snow cover may depend on the relative effects of temperature and precipitation (Raisanen 2008) with regional characteristics.

However, available snow cover datasets have limited this study. For example, in situ snow cover datasets have limited spatial coverage and the snow courses are often specific to their surrounding environment (Dyer and Mote 2006), with more scarce or no observations in the circumpolar regions. Prior to 2000, satellite-based snow cover maps produced by National Oceanic and Atmospheric Administration (NOAA) offer at best 25 km spatial resolution (Armstrong and Brodzil 2002) with daily temporal resolution only available after 1997 (Ramsay 1998). Wang et al. (2005) reported that the NOAA weekly dataset consistently overestimated snow cover extent during the spring melt period, with delays of up to 4 weeks in melt onset. Hence these snow datasets may not be sufficient for identifying spring snowmelt date changes.

We use recently produced daily 5 km APP snow cover maps (Zhao and Fernandes 2009) from 1982 to 2004. Both because they show good agreement with in situ sites and because their almost continuous spatial and temporal coverage makes them especially suitable for analysis of spring snowmelt patterns and snow albedo feedback in association with variability of circumpolar climate ecological parameters (Zhao and Fernandes 2009).

Snow cover data, because of its binary nature, has apparent interannual variability only in transient regions (Zhao and Moore 2006). As a result, climate signals can be only identified in the marginal areas (Clark et al. 1999). As a result we will use spring snowmelt date (Smtd) to investigate snow variability. Quantifying trends in Smtd also provides information relevant to ecosystem studies such as caribou habitat assessment and diagnosing impacts of climate change on permafrost.

Interannual variability in snow cover extent and snowmelt date is tightly correlated to local temperature and precipitation through controls of large-scale atmospheric circulations (Clark et al. 1999; Saito et al. 2004). Understanding the atmosphere-snow interaction is important for better understanding snow albedo feedback and its control factors. Atmospheric circulation teleconnection patterns characterize low frequency climate variability on regional or even hemispheric

scales. The Arctic Oscillation (AO) and North Atlantic Oscillation (NAO) are two of the most important patterns characterizing the NH extratropical climate variability, especially during winter (Hurrell 1995; Thompson and Wallace 1998), and exert influence on spring and summer climate variability (Bamzai 2003; Bojariu and Gimeno 2003; Ogi et al. 2003; Zhao and Moore 2006). Over the North Pacific and North America regions, the North Pacific Oscillation (NPO) and the Pacific/North American pattern (PNA) are two most important modes, influencing North America spring to summer climate. The positive phase of NPO is associated with above-average surface temperatures over the eastern North Pacific, and below-average temperatures over the central North Pacific and eastern North America (Bell and Janowiak 1995). With strong linkage to the strength and location of the East Asian jet stream, both NPO and PNA exert influence on spring snow variability in Canada and the western United States.

The objectives of this paper are to investigate the variability of Smtd over the NH lands using the new APP snow cover product, and driving effects of large-scale atmospheric circulations on the interannual variability in Smtd. Finally a new approach to estimate the temperature sensitive regions (TSRs) of snow cover on the base of seasonal change in temperature (Hall and Qu 2006), is introduced, and the results are compared with the results of Groisman et al. (1994b) which was based on anomalous year composites. With these new TSRs, we also try to explain why there is less contribution in interannual variability of snow factor to snow albedo feedback.

8.2 Data and Methods

The new snow cover dataset, based on a new algorithm (Fernandes and Zhao 2008) applied to the APP daily 5 km Equal Area Scalable Extent (EASE)-Grid composites (Fowler et al. 2000), covers the NH polar region bounded by upper left corner (30°N, 135°W) and lower right corner (30°N, 45°E) between 1982 through 2004 (Zhao and Fernandes 2009).

The APP snow cover products were validated using quality controlled snow depth measurements on a daily basis between 1982 and 2004 over 67 standardized in situ sites in Canada and over 260 station sites in the northern Eurasia. Smtd was defined as the date (on which snow disappears) from which each pixel has no snow at least for three continuous days during the melting period from Day Of Year (DOY) 90 to 244 (corresponding April 1 to August 31) (Zhao and Fernandes 2009).

The monthly NCEP reanalysis (Kalnay et al. 1996) surface air temperature, sea level pressure (SLP), and surface wind fields were used to explore associated atmospheric circulations to the Smtd over the period of 1982–2004. The precipitation dataset is from the Global Precipitation Climatology Project (Huffman et al. 1997), which is a blend of surface and satellite observations that is available since 1979. The ERA40 reanalysis monthly air temperature (see online at <http://www.ecmwf.int/research/era/>) field at 2.5° resolution was only used to calculate TSRs. The reason to estimate TSRs by using ERA40 rather than NCEP

is because a parallel study on quantifying observational snow albedo feedback parameters used the dataset (Fernandes et al. 2009; Qu and Hall 2007).

The AO index (AOI), defined as the normalized difference in zonal-averaged SLP anomalies between 35°N and 65°N, is from (Li and Wang 2003). The NAO index (NAOI), NPO index (NPOI, originally named EP/NP oscillation) and PNA index (PNAI) are downloaded from Climate Prediction Center (CPC) websites, available at <http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml>.

The melting season is defined as a period of three months from April to June (AMJ), the winter from January to March (JFM), and the spring from March to May (MAM), except especially mentioned.

Rather than to estimate TSRs by choosing anomalous years (Groisman et al. 1994b), we estimated TSRs by using seasonal difference of snow fraction divided by temperature difference from April to May (Hall and Qu 2006). For this purpose, the APP snow cover was firstly sub-sampled to 25 km resolution from the original 5 km data and then snow fraction was estimated on the base of an approximate 2.5° grid cell box to match the reanalysis spatial resolution. At the same time, the all sky APP extended daily 25 km albedo product (Wang and Key 2005) was applied. The albedo is averaged over the same 2.5° grid cell box. Finally, the daily snow fraction and daily albedo was averaged temporally to obtain monthly mean fields.

Instead of using the standard deviation to calculate interannual variability, a robust approach has been designed to calculate the variation coefficient in Smtd, based on a range of 68% percentile values symmetric about the median. This new approach can capture major interannual variability without contamination of outliers.

The statistical significance of the regressions and trends were assessed through the use of the Student's *t*-test. When testing significance of a trend in a time series, we used a reduced effective sample size, which is a function of the lag-1 (year) autocorrelation of the regression residuals, in the calculation of the standard error in the trend and in the application of the *t*-test (Angell 1999; Santer et al. 2000). For scalar fields, the regions where the null hypothesis can be rejected at the 90% confidence level are indicated. With regard to wind field, regions where at least one of its components is significant at the 90% level are indicated.

8.3 Results

8.3.1 Interannual Variability in Smtd

Figure 8.1a shows spatial distribution of the mean Smtd during 1982–2004 at 5 km resolution over the NH. Smtd generally increases with latitude consistent with the seasonal march of solar radiation during spring and early summer in the NH. High altitude areas such as the western cordilleras and regions mapped as having permanent ice or snow cover represent clear outliers in this pattern. The approximately zonally distributed Smtd belts also show earlier melt dates at the western margins of

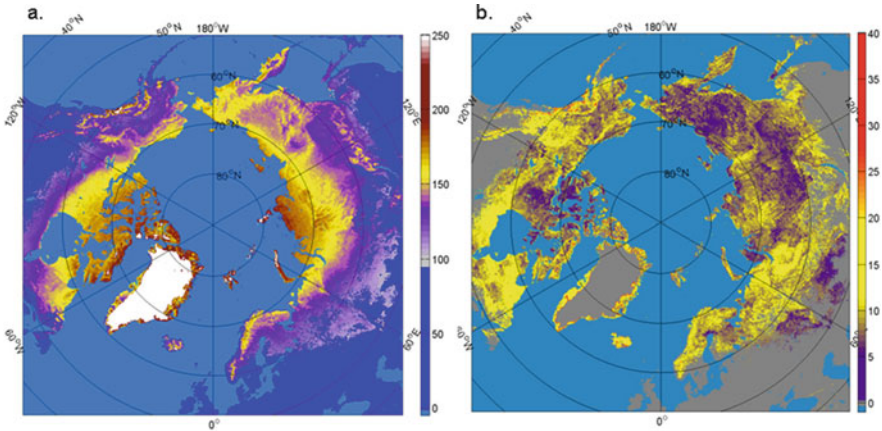


Fig. 8.1 (a) The spatial distribution of the mean Snowmelt dates (Smt_d, unit: DOY) for 1982–2003 over Northern Hemisphere APP domain. The value of 250 (white color) is assigned to permanent snow. (b) The standard deviations of the Smt_d at each pixel over the NH APP domain. Grey colors indicate permanent snow or ice in Greenland and high latitudes or areas where Smt_d is earlier than 90 DOY

continents consistent with warm Pacific and Atlantic currents. Over Eastern Siberia and over Kamchatka peninsula the Smt_d shows clearly topographical effects associated with delayed melt dates over mountains comparing with its basin areas. The latest melting occurs over Northeast Canada and the northern central Siberia.

Figure 8.1b shows the spatial distribution of interannual variability of Smt_d. The higher values with large scale characteristics are mainly located over the western Siberia and northern Europe, the central North Siberian mountains, as well as the eastern Canada and the northwest North America, although some small regions with extreme values are seen over margin areas such as over the western cordilleras, north most Canada, circum-Greenland, Iceland, and islands in Arctic, which are associated with complex surface conditions. There is a latitudinal difference in these high values between East Canada, which is over 50–60°N latitudes, and the central and northwest Eurasia, over 60–70°N latitudes.

Figure 8.2 presents the time series of normalized Smt_d averaged over the two continents, the northern Eurasia (EA, solid lines) and North America (NA, dashed lines), and different latitude bands over the respective continents as well. The mean Smt_d is respectively 140.2 and 148.8 DOY for EA and NA. The two time series of Smt_d (Fig. 8.2a) do not show statistically significant negative trends as expected over the period of 1982–2004 rather with insignificant positive trends, 0.1 day per decade and 0.2 day per decade over EA and NA respectively. The two time series match to some extent in year-to-year variability between the two continents for the period of pre-1998, but diverge since 1998, which may be associated with interdecadal variability in the atmospheric circulations (Zhao and Moore 2009). Figure 8.2b–d show the Smt_d averaged over each latitude band for each continent. Although all show positive trends from 0.1 to 0.4 day per decade except no

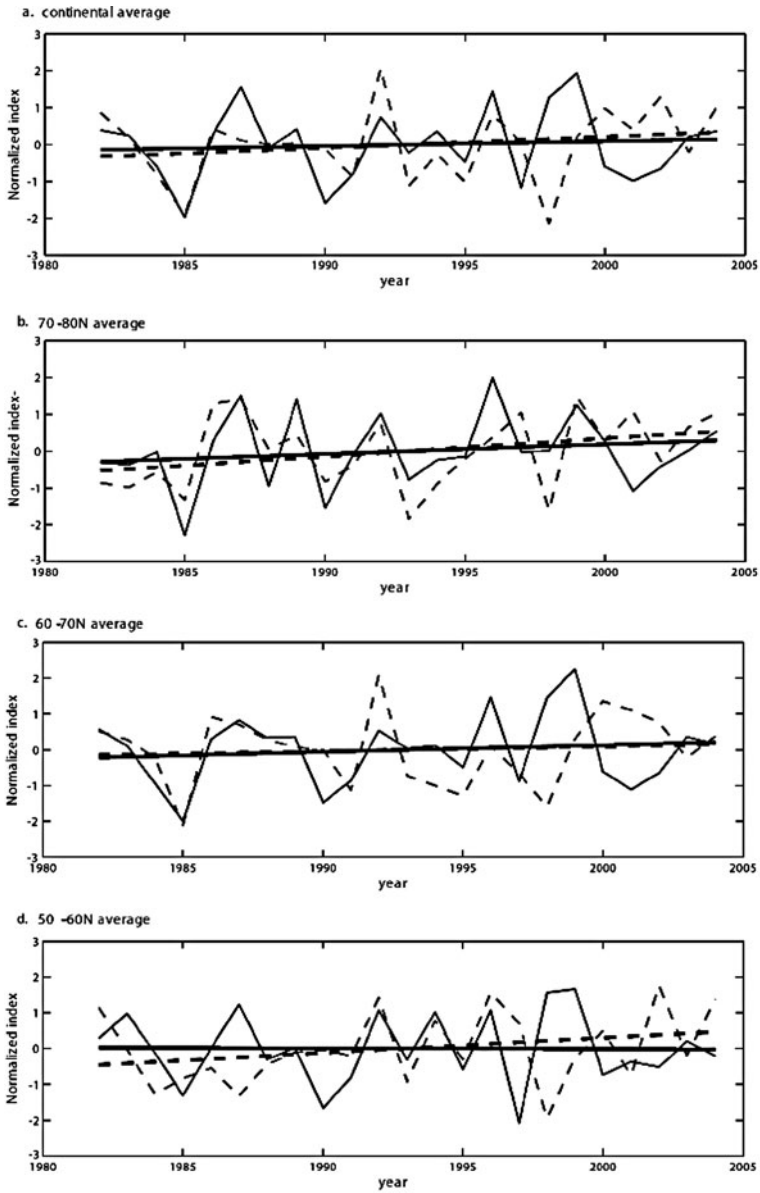


Fig. 8.2 The time series of Smtd averaged over (a) the northern Eurasia (EAmt, solid) and North America (NAmt, dashed). (b) The 70–80°N latitude band. (c) The 60–70°N latitude band. (d) The 50–60°N latitude band. Thick lines are the linear trends over the period of 1982–2004 corresponding EA (solid) and NA (dashed)

trend for Eurasian 50–60°N band, these trends are not statistically significant even at 75% confidence level except the North American 70–80°N band which reaching the 90% confidence level. The divergence post-1998 can also be seen for these latitude bands.

8.3.2 Atmospheric Circulations Drive Large-Scale Interannual Variability in Smtd

To investigate factors driving the Smtd, Fig. 8.3 shows the correlations (Fig. 8.3a and b) and regressions (Fig. 8.3c and d) of the Smtd time series of EA and NA versus the mean melting season (AMJ) surface air temperature field (Fig. 8.3a and b) and the spring mean (MAM) SLP and surface wind fields (Fig. 8.3c and d). From Fig. 8.3a and b, as expected, the negative correlations are dominant over most of area of the continents corresponding to higher Smtd values (delayed melt), confirming that there is a negative snow-temperature interaction mechanism associated with snow albedo feedback. From Fig. 8.3c, the statistically significant pressure anomalies in high latitude, producing pressure gradients between polar region and Siberia, result in statistically significant anomalously northerlies along the continent. This leads to reduced temperatures over this area during melting season (Fig. 8.3a). It can also be seen that the southward extended northerlies over western Siberia correspond to a cold tongue over western Asia (Fig. 8.3a). The negative temperature

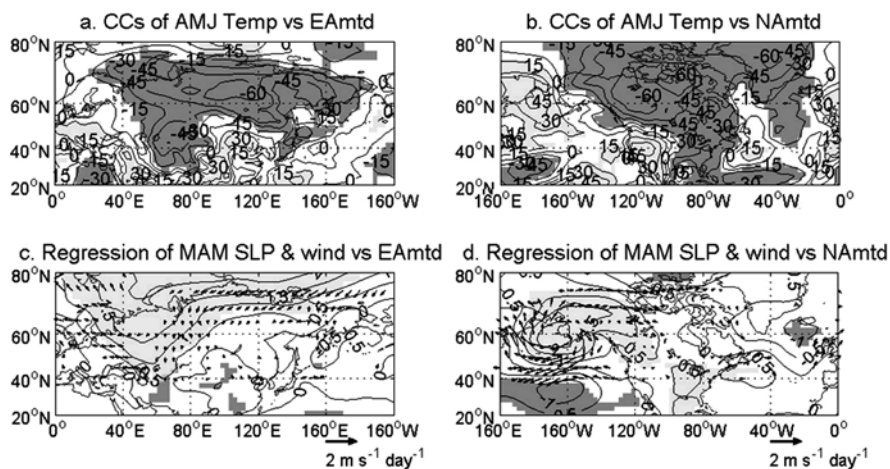


Fig. 8.3 Correlations of (a) EAmt d and (b) NAMt d with AMJ mean temperature field and regressions of (c) EAmt d and (d) NAMt d with SLP and wind fields. Note that in Panels (a) and (b), correlation coefficients have been multiplied by 100, dark (light) shading areas are values larger (less) than 15 (–15), and thresholds for the 90% (95%) confidence level is 0.35 (0.42). In Panels (c) and (d), shaded areas are statistically significant at 90% level for SLP while wind field is only shown at those locations where at least one of its components is significant at 90% level using Student *t*-test

correlation regions over NA (Fig. 8.3c) are related to southward transport of cold air resulting from an anomalously pressure gradient between central NA and eastern NA that seems to be associated with a NPO-like pattern.

The correlations of the EA and NA Smtd with winter mean precipitation field were also calculated (not shown). Specifically, the zero correlations over Eurasia fall within a narrow band extending from northern Europe to northern Japan. North of this line are positive correlations while south of this line are negative correlations, indicating reverse effects of precipitation on Smtd. Over most of areas north of 50°N in NA, there are positive correlations between the Smtd and the winter precipitations (and these positive correlations can extend from winter to spring). Negative correlations are observed southwards of 50°N. Although there are small regions statistically significant at 90% confidence level over both continents, the large continental scale distribution of north–south seesaw structures with positive correlation in higher latitudes and negative correlations in lower latitudes, suggesting potential increase in precipitation under recent warming in Arctic may have cancelled to some extent the effects of increase in spring temperature on the expected negative trend in Smtd (Min et al. 2008).

To further investigate impacts of preceding and instantaneous atmospheric circulations on the spring snowmelt dates, correlations between the monthly AOI, NAOI, NPOI and PNAI from January to June and the various Smtd indices (presented in Fig. 8.2) are calculated and only correlation coefficients in a group that has at least one significant at 95% confidence level for a monthly teleconnection index are presented in Table 8.1. The AOI in March is negatively correlated to EA Smtd including all latitude bands over EA, indicating that a positive AO phase in spring will result in earlier Smtd over EA and vice versa. This is consistent with results from Figs. 8.3a and 8.4c and can be interpreted by comparison with Fig. 8.4a, which shows positive temperature anomalies over EA during melting season associated with a positive AO phase, leading to earlier melt of snow. Similarly, the AOI in April is negatively

Table 8.1 Correlation coefficients (CCs) of Smtd with teleconnection indices. Note that only those CCs in a group with at least one significant at 95% confidence level for a monthly teleconnection index are presented while those under 90% confidence level are not shown

	EAmt	50–60°N	60–70°N	70–80°N	NAmt	50–60°N	60–70°N	70–80°N
AOI (Mar)	–0.43*	–0.45*	–0.35	–0.47*				
AOI (Apr)					0.43*			
NAOI (Feb)	–0.35	–0.46*						
NAOI (May)				0.43*	0.53**		0.67***	0.56**
NPO (Mar)						0.46*		
NPO (Apr)						0.57**		
PNA (Mar)		0.40				–0.43*		
PNA (Apr)						–0.52**		

Note that the table only shows those correlations significant at least at 90% confidence level for a monthly teleconnection index that has at least a correlation reaching the 95% confidence level. The “*” indicates significant at 95% confidence level, the “**” at 99% level, and the “***” at 99.9% level using the Student’s *t*-test.

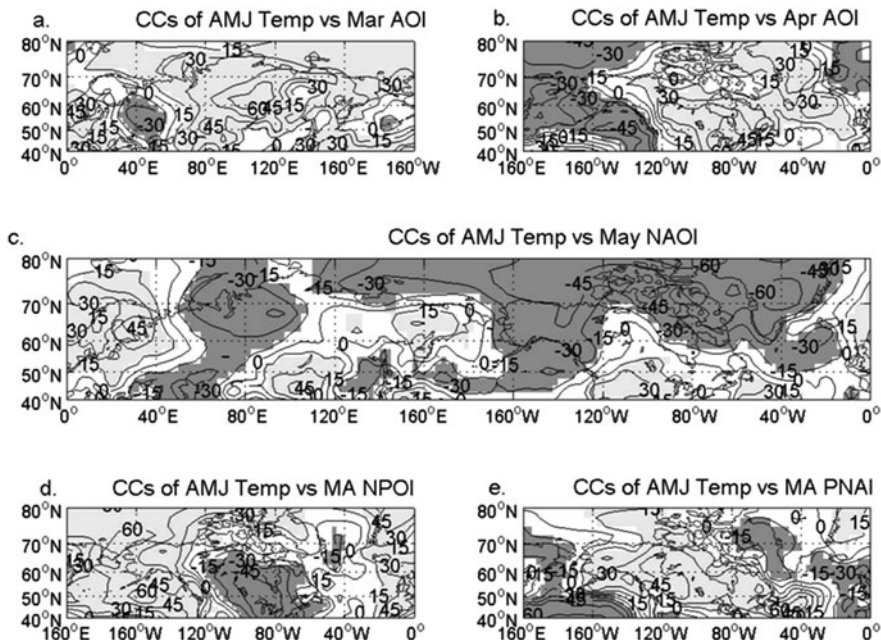


Fig. 8.4 Same as Figs. 8.3a and b but correlations of (a) March AOI, (b) Apr AOI, (c) May NAOI, (d) MA mean NPOI, and (e) MA mean PNAI with AMJ mean temperature field. Note that the spatial domain in Panel c combines those in Panels (a) and (b)

correlated with NA 50–60°N Smtd, where there exists higher interannual variability (refer to Fig. 8.1b), which can be interpreted by comparison between Figs. 8.3b and 8.4b. The delay with a month in correlations with AOI over NA in comparison with over EA may be attributed to a later melt dates over NA as mentioned earlier. The NAOI in May is positively correlated to EA Smtd and NA Smtd in higher latitude bands over both continents (Table 8.1). This linkage can be seen more clearly in Fig. 8.4c where the May NAOI negatively correlated to high-latitude temperatures over these continents inducing a delayed Smtd. It is also noted that the February NAOI is significantly correlated to EA 50–60°N Smtd with a month earlier comparing with AOI (Table 8.1). As consistent with the robust signal over North Pacific in Fig. 8.3d, the NPO in spring (March to April) is positively correlated to NA Smtd over 50–60°N, where there is large interannual variability, while PNA in the same season is negatively correlated with the Smtd. This difference indicates reverse impacts on the East Canada Smtd in comparison with NPO through driving temperature anomalies over NA (Fig. 8.4).

It is speculated that contribution of the preceding and instantaneous atmospheric circulation anomalies to the snow-temperature interaction sensitivity regions will propagate interannual variability into the snow components in the snow albedo feedback (Groisman et al. 1994a; Qu and Hall 2007).

8.3.3 Snow Temperature Sensitivity Regions (TSRs)

Groisman et al. (1994b) presented the relationship between snow cover and surface air temperature by defining the temperature sensitivity regions (TSRs) using 1972–1992 NOAA weekly snow cover dataset. However, the surface air temperature anomalous years chosen for warm (cold) groups are unevenly distributed in 1970s (1980s) respectively. If there were an artificial trend in the satellite observed snow cover time series, say with different era sensors, the TSRs would mainly represent this effect. Hence, following Hall and Qu’s work (2006), based on 18-year mean April and May snow and temperature observations, the seasonal snow temperature sensitivity regions are given in Fig. 8.5a. The spatial pattern of this new TSRs over NA (Fig. 8.5a) is very similar to that in Groisman et al. (1994b). But, over southwest EA areas, the new approach leads to a more continuous distribution in sensitivity regions in comparison with their distribution sensitive to the selection of anomalous years. This indicates a potential improvement of seasonal snow TSRs derived here in comparison with TSRs derived in Groisman et al. (1994b), possibly due to improved snow cover information. The maximum values in the TSRs are located over 55–65°N which is consistent with the significant correlations of Smtd with NCEP temperatures over both continents during snow melting season (Fig. 8.3).

From Hall and Qu (2006), the snow albedo feedback (SAF) parameter can be simply decomposed into components, $SAF = k_3 + k_2 \cdot k_1$, where k_3 is the metamorphosis of the snowpack under a higher temperature, k_1 is the TSRs, and k_2 is the sensitivity of surface albedo to snow cover reduction, a approximate constant in the snow cover component (Fernandes et al. 2009; Groisman et al. 1994a; Qu and Hall 2007). Fernandes et al. (2009) recently diagnosed the SAF components using the snow cover data set presented here together with surface albedo estimates between 1982 and 1999. Table 8.2 gives the values for the SAF components, averaged over

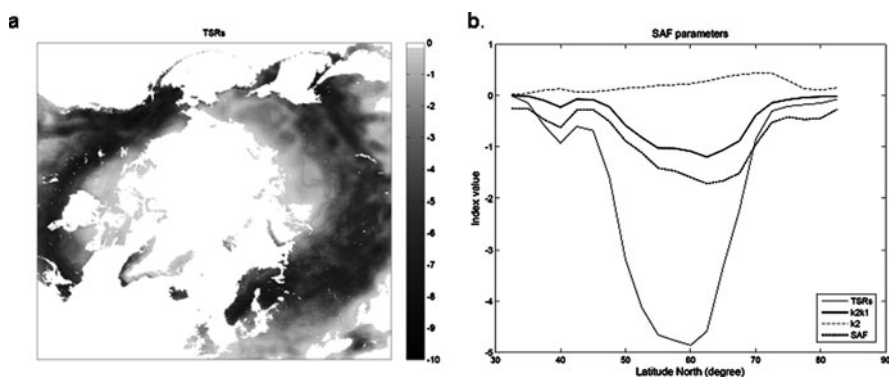


Fig. 8.5 (a) Seasonal snow Temperature Sensitive Tegions (TSRs) based on April to May snow cover fraction and ERA40 monthly temperature at each grid point within APP spatial domain. (b) Latitude dependent SAF parameter (SAF) and snow component (k_2k_1), as well as snow factors (k_2 , k_1), where k_1 is averaged TSRs over latitude bands

Table 8.2 SAF parameters averaged over 1982–2004 and APP domain as well as interannual variability estimated by two times of standard deviation

	$k1$	$k2$	$k3$	$k4$	$k2\ k1$
Mean value	-3.07	0.26	-0.64	-1.26	-0.73
Interannual variability	0.21	0.01	0.05	0.06	0.07

continental and NH domains, and their interannual variability corresponding to the 95% interval for the 18 year mean. Only $k1$ has a relatively large interannual variability, which dominated by interannual variability in atmosphere as analyzed in previous section, while other factors have smaller interannual variability. Due to the small value of $k2$, the interannual variability of $k1$ only contributes to the snow component ($k1 \cdot k2$) with 26 percent. Figure 8.5b clearly shows there exists a latitude difference in major values between $k1$ (55–65°N) and $k2$ (65–75°N), resulting in cancelling not only in values but also in interannual variability in the averaged value of $k1 \cdot k2$.

8.4 Summary and Discussions

In this study, based on the new daily 5 km resolution APP snow dataset, we have produced spring snowmelt dates at each pixel over APP domain for each year over 1982–2004 and observed an expected mean latitude-dependent patterns in spring snowmelt dates together with areas where orographic effects are expected to delay snowmelt. The large interannual variability in snowmelt date together with the absence of obvious trends would be counterintuitive if one hypothesized that the long-term spring warming alone drive snowmelt. Rather, our data provides evidence in support the more complex hypothesis that winter precipitation and spring temperatures interact, in perhaps opposing directions, in their impacts on snowmelt trends in the NH (Brown 2000; Dyer and Mote 2006; Robinson 2003; Raisanen 2008; Ye et al. 1998). Furthermore, the atmospheric circulation variability seems to dominate the interannual variability in Smt. However the strong interannual variability in snow cover only contributed around more than 20% to the snow component in the snow albedo feedback because there is a latitude difference between the albedo-snow sensitivity regions ($k2$), situating in high latitude circum-polar regions (65–75°N, Fig. 8.5b), and the snow-temperature sensitivity regions (55–65°N, Fig. 8.5a). This latitude difference between the two sensitivity factors (Fig. 8.5d) lead to a milder and smoothed contribution of the snow component (Fig. 8.5c) to total SAF.

As expected, negative correlations of the Smt. with surface air temperature field clarified that earlier (latter) continental Smt. are associated with positive (negative) continental temperature anomalies during melting season, consistent with the positive feedback between spring warming and reduced snow (Groisman et al. 1994a). Our results did not show negative trends for various large-scale spatial domains in agreement with the results in Brown et al. (2007), who also did not

find statistically significant trends in spring snow duration over 1979–2003 in central North Canada from in situ, weekly NOAA and passive microwave snow datasets and have attributed to a cooling period in this area, and the work in Vicente-Serrano et al. (2007) who did not find any trend in the date of snow disappearance over NH between 1988 and 2003 based on satellite passive microwave data. There are seesaw-like structures in the correlations between preceding winter precipitations and the Smtd over both continents. The seesaw-like structures over the both continents suggest that the increased arctic precipitation corresponding recent global warming (Min et al. 2008; Raisanen 2008) may have delayed the continental snow melting in spring, which might cancel, to some extent, the earlier melting effects induced by increased spring temperature.

The NPO and the PNA clearly showed existence of significantly correlations with NA Smtd during spring through the impacts on melting season temperatures. Vicente-Serrano et al. (2007) also identified a negative correlation between May PNA index and the date of snow cover disappearance over southern Canada. The spring AO pattern significantly affects the EA Smtd through advecting warm (cold) air into the continent corresponding to its positive (negative) phase (e.g., Wang et al. 2006). These results are consistent with those in Schaefer et al. (2004) in that the AO influences snowmelt date in northern Europe and eastern Siberia and almost no influence in North America. The NAO tightly correlated with NA higher latitude Smtd in May due to combined effects of temperature and precipitation induced by NAO anomalies. The negative correlation of winter NAO with the 50–60°N EA Smtd mainly resulted from temperature effects.

Recent studies show that there is nonlinearity in the climate between North Pacific and the North Atlantic that is a function of the sign of the Pacific Decadal Oscillation (PDO) (Zhao and Moore 2009). When the PDO is in its positive phase, the coupling in the SLP field between the North Pacific and the North Atlantic is stronger resulting in a leading mode of hemispheric SLP variability that resembles the AO pattern. In contrast, when the PDO is in its negative phase, the SLP field in the North Pacific is weakly coupled with that in the North Atlantic resulting in a leading mode of regional scale SLP variability that resembles the NAO. The temporal trend in the co-variation in continental 60–80°N zonal average Smtd (Fig. 8.2) may reflect the existence of the warm PDO phase that began in 1977 followed, after 1998, by a possible reversal to a cool PDO stage (Biondi et al. 2001).

The absence of statistically significant temporal trends in Smtd may be due to both the relative short record and also the complex influence on Smtd during the late 20th century: increased winter precipitation in high latitude (Min et al. 2008), increased spring temperature in Arctic (Wang and Key 2003) and neutral or reversal atmospheric circulations conditions (Overland and Wang 2005). At the same time we noted that the APP snow cover improved the spatial delineation of temperature sensitive regions of snow cover over southwest Eurasia.

In addition to snow variability we also investigated the relationship between temperature sensitive regions, and more generally surface albedo feedback components, with climate indices. The effects of the preceding and instantaneous atmospheric circulation anomalies on the spring snow variability have contributed strong interannual variability to the temperature sensitivity regions. However,

sensitivity to interannual or decadal climate trends is substantially reduced when considering surface albedo feedback components. This suggests that while Smt patterns and trends may be useful indicators of climate change and variability it may be preferable to use diagnostics such as surface albedo feedback components to assess the performance of climate models.

Acknowledgments The authors thank Dr. Y. Zhang, Dr. J. Li, and R. Brown for their pre-reviewing this manuscript. Thanks are also directed to Dr. X. Yang and another anonymous reviewer for their comments and suggestions. This research is supported by Climate Change Program, Natural Resources Canada.

References

- Angell JK (1999) Comparison of surface and tropospheric temperature trends estimated from a 63-station radiosonde network, 1958–1998. *Geophys Res Lett* 26: 2761–2764
- Armstrong R, Brodzik MJ (2002) Northern hemisphere EASE-grid weekly snow cover and sea ice extent version 2, National Snow and Ice Data Center, Boulder, CO, CD-ROM
- Bamzai AS (2003) Relationship between Snow Cover Variability and Arctic Oscillation Index on a Hierarchy of Time Scales. *Int J Climatol* 23:131–142
- Bojariu R, Gimeno L (2003) The role of snow cover fluctuations in multiannual nao persistence. *Geophys Res Lett* 30:L015651
- Bell G, Janowiak J (1995). Atmospheric circulation associated with the Midwest floods of 1993. *Bull Am Meteorol Soc* 76:681–695
- Biondi F, Gershunov A, Cayan DR (2001) North Pacific decadal climate variability since 1661. *J Climate* 14:5–10
- Brown RD (2000) Northern hemisphere snow cover variability and change, 1915–1997, *J Climate* 13:2339–2355
- Brown R, Derksen C, Wang L (2007) Assessment of spring snow cover duration variability over northern Canada from satellite datasets. *Remote Sens Environ* 111:367–381. doi:10.1016/j.rse.2006.09.035
- Clark MP, Serreze MC, Robinson DA (1999) Atmospheric controls on Eurasian snow extent. *Int J Climatol* 19:27–40
- Déry SJ, Brown RD (2007) Recent Northern Hemisphere snow cover extent trends and implications for the snow-albedo feedback. *Geophys Res Lett* doi:10.1029/2007GL031474
- Dyer JL, Mote TL (2006) Spatial variability and trends in observed snow depth over North America. *Geophys Res Lett*. doi:10.1029/2006GL027258
- Fernandes R, Zhao H (2008) Mapping daily snow cover extent over land surfaces using NOAA AVHRR imagery. In: Proceedings of the 5th EARSeL Workshop: Remote Sensing of Land Ice and Snow, Bern, 11–13 February
- Fernandes R, Zhao H, Wang X et al (2009) Controls on Northern Hemisphere snow albedo feedback quantified using satellite earth observations. *Geophys Res Lett*. doi:10.1029/2009GL040057
- Fowler C, Maslanik J, Haran T et al (2000) AVHRR Polar Pathfinder twice-daily 5 km EASE-grid composites V003. Boulder, Colorado USA, National Snow and Ice Data Center. Digital media
- Groisman PY, Karl TR, Knight RW (1994a) Observed impact of snow cover on heat balance and the rise of continental spring temperatures. *Science* 263:198–200
- Groisman PY, Karl TR, Knight RW et al (1994b) Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere. *J Climate* 15:1633–1656
- Hall A, Qu X (2006) Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys Res Lett*. doi:10.1029/2005GL025127
- Hall A, Qu X, Neelin JD (2008) Improving predictions of summer climate change in the United States, *Geophys Res Lett*. doi:10.1029/2007GL032012

- Huffman GJ, Adler RF, Arkin P et al (1997) The global precipitation climatology project (GPCP) combined precipitation dataset. *Bull Am Meteorol Soc* 78:5–20
- Hurrell JW (1995) Decadal trend in the North Atlantic Oscillation: regional temperatures and precipitation. *Science* 269:676–679
- Kalnay E, Kanamitsu M, Kistler R et al (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–471
- Li J, Wang J (2003) A modified zonal index and its physical sense. *Geophys Res Lett.* doi:10.1029/2003GL017441
- Massom R (1991) Satellite remote sensing of polar regions: applications, limitations and data availability. Belhaven Press and Lewis Publishers, London
- Min SK, Zhang X, Zwiers F (2008) Human-induced Arctic Moistening. *Science* 320:518–520
- Ogi M, Tachibana Y, Yamazaki K (2003) Impact of the Wintertime North Atlantic Oscillation (NAO) on the summertime atmospheric circulation. *Geophys Res Lett.* doi:10.1029/2003GL017280
- Overland JE, Wang M (2005) The third Arctic climate pattern: 1930s and early 2000s. *Geophys Res Lett.* doi:10.1029/2005GL024254
- Qu X, Hall A (2005) Surface contribution to planetary albedo variability in cryosphere regions. *J Climate* 18:5239–5252
- Qu X, Hall A (2007) What controls the strength of snow-albedo feedback? *J Climate* 20:3971–3981
- Raisanen J (2008) Warmer climate: less or more snow? *Climate Dynam.* doi:10.1007/s00382-007-0289-y
- Ramsay B (1998) The interactive multisensor snow and ice mapping system. *Hydrol Processes* 12:1537–1546
- Robinson DA (2003) Recent variability of Northern Hemisphere snow cover. paper presented at Seventh Conference on Polar Meteorology, Hyannis, Massachusetts, USA
- Saito K, Yasunari T, Cohen J (2004) Changes in the sub-decadal covariability between Northern Hemisphere snow cover and the general circulation of the atmosphere. *Int J Climatol* 24:33–44
- Santer BD, Wigley TML, Boyle JS et al (2000) Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J Geophys Res* 105:7337–7356
- Schaefer K, Denning S, Lonard O (2004) The winter Arctic Oscillation and the timing of snowmelt in Europe. *Geophys Res Lett.* doi:10.1029/2004GL021035
- Thompson DW, Wallace JM (1998) The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys Res Lett* 25:1297–1300
- Vicente-Serrano SM, Grippa M, Le Toan T et al (2007) Role of atmospheric circulation with respect to the interannual variability in the date of snow cover disappearance over northern latitudes between 1988 and 2003. *J Geophys Res.* doi:10.1029/2005JD006571
- Wang D, Wang C, Yang X et al (2006) Winter Northern Hemisphere surface air temperature variability associated with the Arctic Oscillation and North Atlantic Oscillation. *Geophys Res Lett.* doi:10.1029/2005GL022952
- Wang L, Sharp M, Brown R et al (2005) Evaluation of spring snow covered area depletion in the Canadian Arctic from NOAA snow charts. *Remote Sens Environ* 95:453–463
- Wang X, Key JR (2003) Recent trend in Arctic surface, cloud, and radiation properties from space. *Science* 299:1725–1728
- Wang X, Key JR (2005) Arctic surface, cloud, and radiation properties based on the AVHRR Polar Pathfinder dataset. Part II: recent trends. *J Climate* 18:2575–2593
- Ye H, Cho H-R, Gustafson PE (1998) The changes in Russian Winter Snow accumulation during 1936–83 and its spatial patterns. *J Climate* 11:856–863
- Zhao H, Moore GWK (2006) A seasonally lagged signal of the North Atlantic Oscillation (NAO) in the North Pacific. *Int J Climatol* 26:957–970
- Zhao H, Fernandes R (2009) Daily snow cover estimation from AVHRR Polar Pathfinder data over Northern Hemisphere land surfaces during 1982–2004. *J Geophys Res.* doi:10.1029/2008JD011272
- Zhao H, Moore GWK (2009) Temporal variability in the expression of the Arctic Oscillation in the North Pacific. *J Climate* 22:3110–3126

Chapter 9

MODIS Snow Monitoring Over the Tibetan Plateau

Li Xu

Abstract Permanent and seasonal snowpack over the Tibetan Plateau (TP) is an important component of earth climate system. Due to the limitation of conventional observations and previous satellite image, however, the snow cover variability over the TP has been a challenging problem for climate study. High-resolution (500 m) snow mapping obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) provided an excellent opportunity to define details in the spatial and temporal snow cover distribution and variability in this region. In previous evaluation, the accuracy of MODIS/Terra snow cover data is evaluated by comparing the data with in situ Chinese station snow observations and show a overall 90% accuracy over the TP. Statistical analysis is then performed to explore the distribution and variability of snow cover based on the available data from 2000 to 2006. The most persistent snow cover is located in the southern and western edges of the TP, especially major mountain ridges and the western Yarlung Zangbo Valley. The higher snow cover fractions (SCFs) are mainly concentrated in the regions where the elevation is higher than 6000 m. The snow persistence varies in different elevation ranges and generally becomes longer with increases in the terrain elevation. In addition, the spatial distribution of the snow cover not only depends on the elevation, but also varies with the terrain features (e.g., aspects and slopes) in the local areas. With 7-year observational data, seasonal and interannual variability of snow cover has been detected, but the persistence of the seasonal and interannual anomalies in snow cover is generally lacking due to the small sample of data available. There are slight decreasing trends in SFCs over the TP during 2000–2006.

L. Xu (✉)
Department of Atmospheric, Oceanic and Earth Science, George Mason
University, Fairfax, VA 22030, USA
e-mail: lxu2@gmu.edu

9.1 Introduction

The Tibetan Plateau (TP), the largest highland plateau over the world, locates at the southwest of China with an average elevation more than 4000 m above sea level. Due to its high elevation, the TP is often called “the roof of the world.” Seasonal snow cover over the Tibetan Plateau (TP) is a unique feature in global snow maps compared with same latitude other regions. With the highest mountains in the world, snow cover can persist during all seasons over the high altitudes in the TP. Previous studies have demonstrated that snow cover over the TP has a strong impact and forcing with general circulation and Asian monsoon systems during spring and summer (Bamzai and Shukla 1999; Wu et al. 1995; Yong 1995). In addition, snow cover over the TP is a vital water source for eastern and southern Asia. The largest rivers of in the nearby region, such as the Yangtze River, Yellow River and Yalung Zangbu river, have their headwaters there. Furthermore, snow cover is a comprehensive indicator of the mean conditions of temperature and precipitation over the TP and its surrounding areas. In the global warming background, snow cover variability over the TP is an important problem in regional climate studies (Qin et al. 2006).

However, accurate observations or monitoring of the snowpack over this region are difficult to obtain. Before the era of “satellite meteorology,” traditional snow observations were usually obtained from ground-based meteorological networks. In these meteorological stations, only the presence or absence of snow along with snow depth is measured on a daily basis by the snow stick. At present, there are only 115 conventional ground-based stations managed by the Chinese government over the TP. Most of these stations are located in the inhabited, lower-altitude river valleys, where elevations are usually below 4500 m. Due to the scarcity of ground-based, in situ stations, it is difficult to adequately capture the spatial variability of snow cover, particularly in remote, difficult to access regions such as the TP.

With advances in remote sensing technology, satellite-derived snow cover information has become an important alternative data source. Weekly snow mapping of the Northern Hemisphere using National Oceanographic and Atmospheric Administration (NOAA) National Environmental Satellite, Data and Information Service (NESDIS) data began in 1966. This snow-cover data is one of the longest and longest satellite record of this important climatological variable, and it has been studied intensively by Frei and Robinson (1998). Guo et al. compared the NESDIS weekly snow product with ground snow depth observations over China in past the 30 years. They found significant differences between the NESDIS snow mapping and in situ Chinese ground-based observations and concluded that the NESDIS snow cover products tend to overestimate seriously snow cover over the TP. The cause of the overestimation is mainly due to the coarse spatial resolution (89×89 grids over Northern Hemisphere in the NESDIS data, wherein the data do not represent well the detailed information of patchy and shallow snow cover over the TP and, instead, just call such areas completely snow covered. Based on the passive microwave remote sensing, Che et al. (2003) point out the serious overestimate the snow water equivalent (SWE) over the TP and must be corrected over this region. Ice core record also could provide snow anomalies in the some points (Thompson et al. 1989), but its information is very limited.

With improving in the polar-orbit satellite and remote sensing technology, the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Terra satellite was launched with a complement of five instruments, one of which is Moderate Resolution Imaging Spectroradiometer (MODIS) On December 18 1999. Besides the comprehensive observations on cloud, ocean, and earth surface characteristics available from the Terra MODIS, the snow-cover product has been available since February of 2000. With substantially improved spatial resolution (500 m globally), high temporal frequency (daily), enhanced capability to separate snow and clouds (Hall et al. 2001, 2002) due to more spectral bands (particularly in the short-wave infrared), as well as a consistently applied, automated snow-mapping algorithm (Riggs and Hall 2002), MODIS provides an excellent opportunity to study snow cover over large, relatively inaccessible regions such as the TP. In particular, the single satellite platform provide excellent consistence with MODIS snow data that are hardly to maintain in the previous dataset.

9.2 Data

9.2.1 MODIS Snow Mapping

A variety of snow and ice products are available at a variety of spatial and temporal resolutions (Hall et al. 2001). These data range from 500 m resolution, 2330 km swath snow-cover map to climate-modeling grid (CMG) products on a latitude/longitude (cylindrical equidistant projection). The 8-day composition snow mappings in 500 m (MOD10A2-V4) and 0.05° climate modeling grid (CMG) datasets (MOD10C2-V4) are obtained from the National Snow and Ice Data Center (<http://nisdc.org>). Compared with the daily snow cover product MOD10A1-v4 and MOD10C1-v4, the 8-day composition snow mappings greatly reduce the percent of cloud obscuration or masked pixels, from near half to less than 7% over the TP. Furthermore, the 8-day composition can cover the entire TP areas without any missing data due to orbit swaths (Riggs et al. 2000).

According to previous study by Pu and Xu (Pu et al. 2007), an evaluation of the MODIS 8-day composite snow cover mapping product (MOD10A2) using available in situ Chinese snow observations shows the overall accuracy is about 90%. Total error over the TP is about 10% and this compares favorably with previous studies showing global average errors of 8%. The omission error and snow detection rate depend on both the persistence of snow during 8-day periods and the snow depth, i.e., very shallow snow depths (<5 cm) cause omission errors. In the snow season (November–May), the total error is relatively larger compared to the snow deficient season (June–October). The accuracy is highest in the lower elevation, and the total error increases with elevation. Overall, the evaluation results show that the 8-day composite snow cover data are typically quite accurate in terms of the snow detection rate over the TP.

Previous satellite snow products, such as the NOAA snow maps at $2 \times 2^\circ$ resolution and the IMS snow maps at 24 km resolution after 1999, provide only binary

information about presence/absence of snow. Because of the inhomogeneous and patchy properties of snow cover in nature, the binary yes/no information does not fully represent the snow status over the complex terrain of the TP. In practice, snow cover is frequently, if not always, defined by the portion of land covered by snow. Snow cover fraction (SCF) is defined as the portion of the land surface which is “seen” by a satellite or other airborne imaging instrument as being covered with snow and provides a better indicator of snow cover in a grid box. Before remote sensing, however, SCF was difficult to measure directly in situ.

With MODIS high-resolution observations, MOD10C2-v4 products, defined as snow cover fraction (SCF) at 0.05° climate modeling grids (CMG), have been derived from the MOD10A2-v4 dataset. Three parameters are provided by MOD10C2-V4 in each CMG grid, including snow cover fraction, cloud fraction and confidence index (CI).

The snow cover fraction (SCF) or snow percent in each CMG grid is defined as

$$f_{\text{snow}} = \frac{n_{\text{snow}}}{n} \times 100$$

where n is the total number of pixels in this CMG grid:

$$n = n_{\text{snow}} + n_{\text{clear_land}} + n_{\text{cloud_land}} + n_{\text{cloud_obscured}} + n_{\text{unknown}}$$

n_{snow} is the total pixels identified as snow in this grid, $n_{\text{clear_land}}$ for snow-free land, $n_{\text{cloud_obscured}}$ for cloud and n_{unknown} for other types of land.

The fraction of the CMG grid obscured by the cloud is computed by the ratio of cloud obscured pixel with total number of pixel. The Confidence Index (CI) is defined as the fraction of all pixels determined to be cloud-free land which is an index to express the confidence of information based on the number of valid observations available in this CMG grid.

$$f_{\text{cloud}} = \frac{n_{\text{cloud_obscured}}}{n} \times 100 \quad \text{CI} = \frac{n_{\text{snow}} + n_{\text{clear_land}}}{n} \times 100$$

As pointed out by Hall et al. (1999), the biggest obstacle in monitoring snow cover by visible imagery is cloud-obscuration. In the MOD10C2-V4, the percent of cloud-obscured pixels has been reduced greatly by compositing 8 days of visible images. Nevertheless, on average roughly 7% of the land is still obscured by clouds. Due to the cloud obscuration, the SCF obtained from MODIS is underestimated. As a reasonable estimation based on the visible portion of the pixels, assuming that the cloud-obscured pixels are all snow-free, there are no underestimations of SCF. If the snow cover percent under cloud obscuration is same as the snow cover percent in the cloud-free area, the underestimation of SCF is equal to $f_{\text{snow}} * (1 - \text{CI})/\text{CI}$. The larger the CI, the smaller the underestimation. At the other extreme, the entire area blocked by cloud may be snow-covered, making the maximum underestimation of SCF equal to the cloud percent.

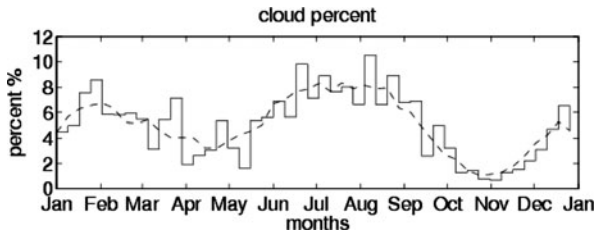


Fig. 9.1 Mean cloud percent over the TP in the MOD10C2-V4 dataset during 2000–2006. *Dash line* is five points moving average
Modified from Xu (2007)

Figure 9.1 shows the average cloud percent over the TP during 2000–2006. The dash line is a 5-point moving average. In the spring and autumn, the cloud obscuration is relatively low (<2%), but the cloud percent is relatively high (>6%) in the summer. The cloud obscuration will cause a slight underestimation of SCF over the TP. As mentioned in the previous chapter, the commission error of MODIS snow maps is larger than omission error over the TP. As a result, the snow algorithm overestimates the snow cover in 500 m pixels. The underestimation due to cloud obscuration partly compensates the overestimation due to commission error. Thus, SFCs from MOD10C2 could be used in climate analysis.

9.2.2 GTOPO30 DEM Data

In this study, the TP is defined as the region between 70–110E, 25–45 N with an elevation higher than 2000 m which covering about 3.6 million square kilometers (detail domain see (Pu et al. 2007)). The USGS GTOPO30 Digital Elevation Model (<http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html>) provides elevation data at 30-arc seconds (approximately 1 kilometer in the Equator). The 30-arc seconds GTOPO30 data is resampled to 0.05° CMG grid by area averaging.

9.3 Seasonal Variations of Snowpack Over the Tibetan Plateau

9.3.1 Snow Distribution and Annual Cycle

The quantitative snow cover distribution over the TP was analyzed based on the SCF observed from Terra satellite during 2000–2006 (Xu 2007). The monthly mean SCF was computed by the weighted averaging of 8-day MOD10C2 data. Extreme spatial variability in monthly snow cover is exhibited due to complex terrain. Specifically, higher SCFs correspond well with the large mountains along western and southern edge. The most persistently snow-covered areas (over 50% of area coverage and 70% of the time) are concentrated in the Himalaya, Kunlun, Karakoram ranges

and the western part of the Yarlung Zangbo Valley, especially in the combined areas of Karakoram and Kunlun mountains which are the most heavily glaciated regions in the world outside of the Polar regions. In the southeastern part of the TP, SCF is also relatively high because the moist air goes up along the Yarlung Zangbo Valley from the southern region. In contrast, due to huge shielding from the Himalaya and Karakoram mountains, most of the interior of the TP has less snow-cover persistence, although the averaged elevation is beyond 4000 m (figure not shown). The distribution of snow cover is strongly linked to the available moisture over the Tibetan plateau. In the southern flank of TP, the available precipitable water comes from the tropical ocean in the south, including the Indian Ocean and the Arabian Sea. The decreasing dramatically from above 10 kg/m^2 to near zero along the Himalayas and Karakoram. Especially, the Pamirs and Tibetan Plateau form a horn-shaped village in the western part of the Himalayas, which further enhances the convergence of moisture and terrain precipitation in the western edge of TP. Due to the huge shield effect by the Himalayas and the Karakoram, there is little moisture in the vast interior of TP. In particular, the precipitable water is near to zero in the center of TP, corresponding to the area with small SCF in the winter. On the other hand, the available moisture in the northern flank of TP is mainly conveyed from remote Atlantic and Arctic oceans by westerly flow, contributing to most of the snowfall in Tianshan region. In the winter, large scale vertical motion over TP descends, which also greatly contributes to the sublimation of snow cover in the interior of TP.

Figure 9.2 shows the annual percentage rate of large SCF over the TP, defined as the percent of days annually in which SCF is larger than 50%. The higher percentage is concentrated in the Himalayas, Kunlun, Karakoram and the western part of Yarlung Zangbo valley, consistent with the larger SCF regions in the winter and spring. Especially in the coalesced region of the Karakoram and the Kunlun

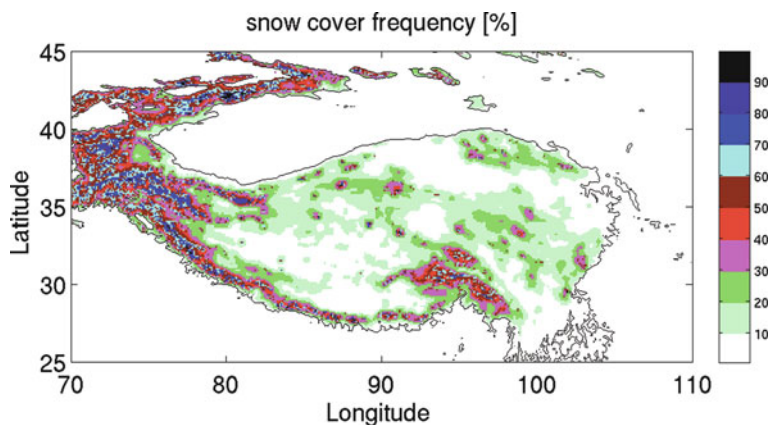


Fig. 9.2 The frequency of heavy snow-covered (SCF > 50%) based on Snow Cover Fraction. The *black line* represents the outline of TP by the contour line of 2000 m
Modified from Pu and Xu (2009)

Mountain, the frequency is larger than 70%. In the southeastern part of TP, the frequency is also relatively large due to the moist airflow that goes up along the Yarlung Zangbo valley from the Indochina Peninsula. Most of the interior TP has a lower frequency of large SCF. Only in the major huge mountain, such as the Nyainqentanglha Mountain, is the frequency relatively higher than the nearby region. In the Qaidam and Tarim basins, the frequency of heavy snow-cover is lowest. In the area of elevation within 2000–4000 m, the median of the duration of more than 50% in SCF is about 22 days, compared to 56 days in region of 4000–6000 m and 195 days in the region of above 6000 m.

The snow cover also displays large intra-seasonal variability. Figure 9.3 shows the annual cycle of SCF and the standard deviation in the different elevation regions, as calculated from available data over the period from February of 2000 to the end of 2006 in 8-day time interval. Strong seasonal variabilities in SCF are found over the whole area of the TP (Fig. 9.3a). From early October to late April, overall SCFs over the TP are greater than 25%, with relatively large standard deviations

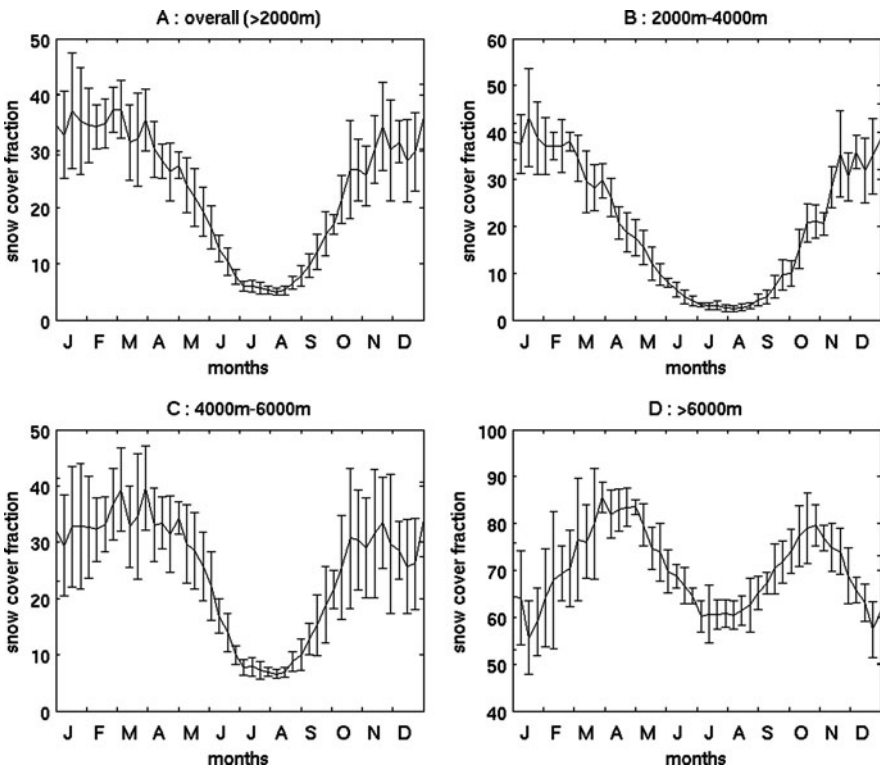


Fig. 9.3 Snow cover fraction (SCF) averaged over the areas in different elevation ranges: (a) Over the whole area of the TP (>2000 m); (b) 2000–4000 m; (c) 4000–6000 m; (d) beyond 6000 m. The curve on each panel shows the averaged SCF of the period of 2000–2006. The error bars show the standard deviations, indicating the variations of snow cover during 2000–2006. Modified from Xu (2007)

reflecting the inter-annual variability from year to year. The SCF, depending upon the elevation, peaks in January at the lower elevations and is delayed as late as in April above 6000 m, and then progressively decreases and reaches minimum values (except above 6000 m) of about 1–5% in the July–August period. The interannual variability is relatively small during the summer months. At elevations above 6000 m (Fig. 9.3d), the average SCFs are greater than about 60% in all seasons. In the area above 4000 m (Fig. 9.3c and d), SCFs show two maximums, one in later fall (November) and another in early spring (March). These two peaks are consistent with the two maximums of annual snowfall corresponding in November and March (figure not shown). This two-peaks characteristic reflects a possible negative (self-regulating) snow cover feedback as suggested by Walland and Simmonds (1996): with sudden SCF increasing after a snow storm in November, the air temperature in the lower troposphere decreases and static stability of atmosphere increases; this reduces the probability of later snowfall; reducing snowfall further results in decreasing SCF by snow blowing event and sublimation in the winter; decreasing SCF in land surface increases the sensible heating to the atmosphere and decreases the static stability, further increasing the probability of snowstorms in March. This negative feedback keeps the snow cover relatively stable over the TP in the winter. Furthermore, snow blowing events by strong westerly winds at upper levels (East Asian Jet Stream (Moore 2004)) in winter also contribute to the reduction of SCF in the higher elevation regions.

9.3.2 Terrain Characteristics of Snow Cover Distribution

Based on the GTOPO30 DEM data, the distribution of SCF over the TP with various terrain features is further analyzed. In geography, slope is defined as the angle of topographic surface to the horizontal plane, and aspect refers to the direction of the incline. Based on the regular grid of DEM, the slope is computed by the maximum change in elevation over distance between each DEM cell with its 8 neighbors. The slope indicates the steepness of terrain, usually shown in the upward angles from the horizon. The aspect identifies the steepest down-slope from each cell to its neighbors, shown in the angles of the surface clockwise from the north. The curvature measures the second derivative of elevation at each cell, akin to the slope of the slope, which is used to describe the convex/concave characteristics of the terrain. A positive curvature indicates that the surface is upwardly convex in the cell and a negative curvature indicates that the surface is upwardly concave. A value of zero indicates that the surface is flat. Matlab function “gradientm” is used to compute the slope and aspect based on a finite-difference approach. In this study, the coarse CMG grid (roughly 5 km) would underestimate the local slope, because the terrain between two adjacent points is modeled as a linear variation, while actual terrain can vary much more abruptly over such a distance.

To study the snow cover distribution in the different aspects, the slopes are classified into four categories: northward/north-facing slopes ($-45/315 < \text{aspect} < 45$), eastward/east-facing slopes ($45 < \text{aspect} < 135$), southward/south-facing slopes ($135 < \text{aspect} < 225$) and westward/west-facing slopes ($225 < \text{aspect} < 315$). In

the distribution map of the aspect (figure not shown), it is easy to recognize that slopes are commonly southward in the southern edge of the plateau. On the other hand, the slope is commonly northward in the northern edge. In the interior of the plateau, the aspects are extremely inhomogeneous.

Figure 9.4 demonstrates the mean SCF variation as a function of slope in the different aspects during four stages, the snow maximum season (February), the ablation season (May), the minimum season (August) and the accumulation season (November). Regardless of the steepness of slopes, the snow cover fraction is larger on the southward and eastward slopes. With increasing slope, the average SCF increases on southward and eastward slopes. In contrast, the SCF keeps a relatively constant value or even a slight decrease in the northward and westward slope. This feature is also a possible result of the uplifting of moisture along the southern and western edge of the TP. In the snow maximum season (February), the SCF does not show much difference in the flat slopes (slope < 5°). With increasing slope, the SCF in southward and eastward slopes grows rapidly.

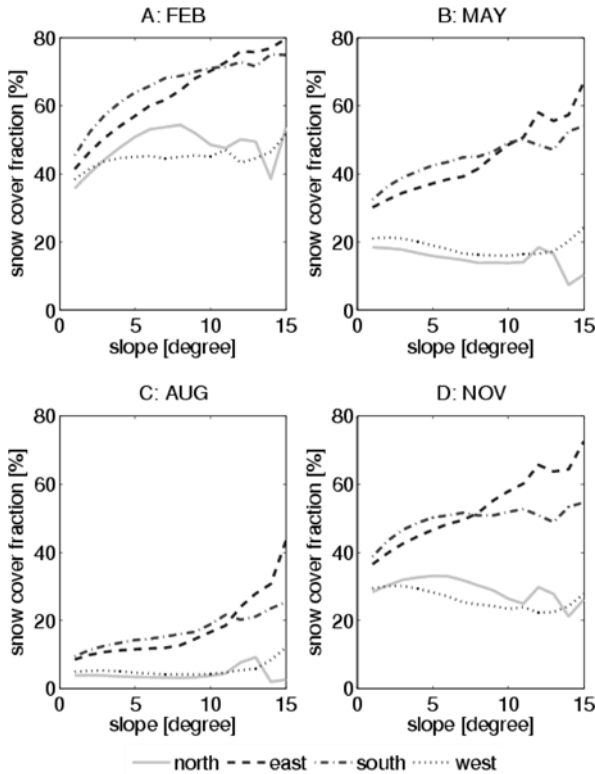


Fig. 9.4 Mean SCF as a function of slope in the different aspects during the snow maximum season (February), the ablation season (May), the minimum season (August) and the accumulation season (November). X-axis is the slope in degree and y-axis is the snow cover fraction in percent. The statistics are based on the Terra snow monitoring from 2000 to 2006 Modified from Pu and Xu (2009)

The average SCF on the eastward slope is larger than the southward slope when the slope is larger than 10° . This feature is more evident in the snow minimum season (Fig. 9.4c). A possible reason is that the southward slope receives more solar radiation than the eastward slope, leading the biggest SCF in the summer located on the steep east-facing mountain. Snow cover in the northward and westward slopes is not sensitive to the slope. In the ablation and accumulation season, the SCF even decreases with the growth of the slope.

Based on the high resolution MODIS snow data, the evolution of snowlines over the TP can be pioneer explored for the first time. In this study, the snowline is defined as the elevation in which snow cover fraction is more than 50% (mostly-snow-covered). Figure 9.5 shows the mean elevation of snowlines on four different aspects.

In winter, the snowline is lowest on northward slopes (roughly 2300 m) during January. Larger snowline differences are shown for the four different directions. The snowline is roughly 3000 m Above Sea Level (ASL) minimum at southward slopes. It is consistent with the common concept that the snowline on the southward slopes is much higher than at the northward slopes. In addition, the snowline on the windward slopes (westward slopes) is also lower than the snowline in the leeward slopes (east slopes). The range of variation of snowlines between different aspects is roughly 1100 m. From winter to summer, the snow cover melts and snowlines rise correspondingly. The snowline in northward slopes rises faster than in other

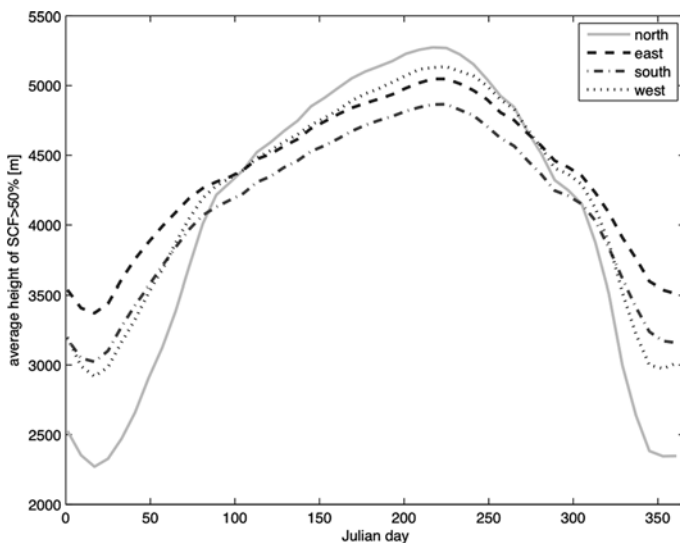


Fig. 9.5 Annual evolution of snowlines in the four different aspects. The x -axis is Julian day and the y -axis is the elevation. The snowline is defined as the elevation in which the SCF is larger than 50%

Modified from Pu and Xu (2009)

directions and finally reaches 5100 m. The snowline in the southward slopes, however, rises much slower and stops at 4700 m ASL maximum. In the meantime, the difference of snowlines in the four aspects reduces to 400 m. The annual variability of snowlines is largest in the northward slopes and smallest in the southward slopes.

Why is the snowline at the southward slopes lower than the northward slopes during the summer? Author speculates that this characteristic is possible as a result of special distribution of huge mountains over the TP. The southward slopes concentrate at the southern edge of the TP where sufficient air moisture and terrain lifting will produce deeper snow cover on the ground than other regions. Although southward slopes receive more solar radiation during summer, the snow albedo effect could still cause a relatively lower snowline. Further field investigations are needed to explore the reason.

9.3.3 Snow Ablation

Snow ablation is the main land surface process during spring that influences the exchange of water and energy with the atmosphere. Figure 9.6 shows the day of onset of snow ablation averaged for different elevations and aspects. The onset date is defined as the Julian day on which the SCF is continuously decreasing for at least three 8-day periods after its peak in the winter.

In lower elevations, the snow ablation begins at the end of February. With increasing elevation, snow ablation is postponed accordingly in the middle of March at 3000 m, April at 4000 m and May above 5000 m. At the same elevation, the snow melts later on the southward and eastward slopes due to deeper snow. Below 5000 m, snow in the northward slopes melts rough 10 days earlier compared to the southward

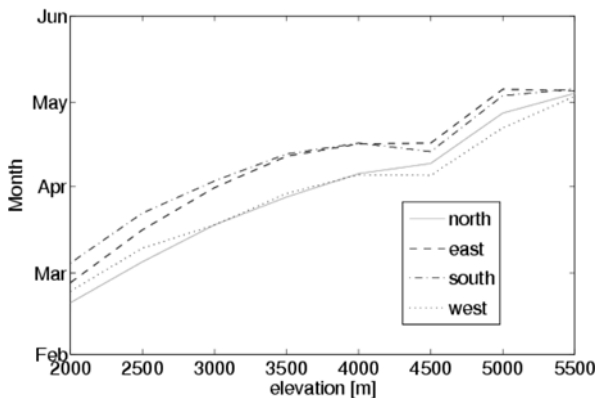


Fig. 9.6 Snow melting (ablation) date as a function of elevation in different aspects. The snow melting date is defined as the Julian day on which the SCF begins continuous decreasing for at least three 8-day periods after its peak in the winter

Modified from Pu and Xu (2009)

slope. The departure in the melting date between the different aspects disappears above 5500 m. From late winter to early spring (Feb-Mar), the snow in the lower elevation regions around the edge of the TP begins to ablate. The SCF in most of the interior of TP, however, is still increased (negative decrement) at that time. The majority of large decrements (decrement $> 10\%$) concentrate in the area between 2000–3000 m (39%) and 3000–4000 m (30%). In early spring (march-April), the snow begins to ablate in the higher elevation regions at Karakoram and Yalung Zhangbo Valley. Simultaneously, the increasing SCF in the interior of TP is significantly reduced. The ratio of large decrement in the area of 4000–5000 m is rapidly increased. After April, the snow is ablated over whole TP except the permanent snowpack and glacier regions. The large ablations concentrate in the area between 3000 and 5000 m. From May to June, more than half of the large decrement occurs in the area above 4000 m.

9.3.4 Interannual Variability and Linear Trend

The reductions of snow cover and glacier due to global warming have been evidenced in the high mountains and the Polar region. Although short in length, the time series of the mean SCF over the whole TP during 2000–2006 were analyzed based on linear regression (Fig. 9.7). The linear trend in the mean SCF, for the whole period, is almost flat over the TP during 2000–2006, although a slight decreasing trend (about -0.34% per year) is found. The minimum SCF in each year also shows a negligible trend of -0.01% per year. The maximum of SCF, however, shows a notable decreasing trend of roughly -1.07% per year. During 2000–2006, the maximum of SCF is roughly reduced 8%.

Although the minimum of SCF over the whole TP does not show a clear decreasing trend, the snow cover fraction over 5000 m does show a notable decreasing

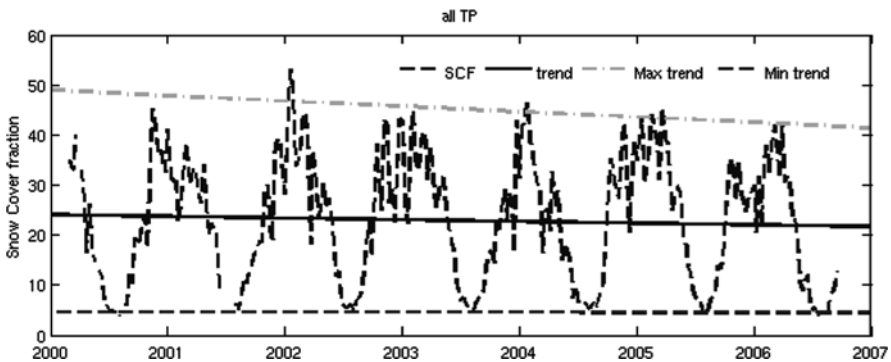


Fig. 9.7 Interannual variability of the average snow cover fraction over the TP from the 2000 to 2006 (dash line). The line shows the linear trend over 7 years period (-0.34% /year). The dash-dot line shows the linear trend of the maximum SCF in the each snow year (-1.07% /year) and the dash line shows the linear trend of the minimum SCF in each year (-0.01% /year)

Modified from Pu and Xu (2009)

trend in the snow minimum season (August) during 2000–2006. The linear decreasing trend is relatively small in the area of 5000–5250 m ($-0.26\%/year$), but increases with the elevation. In the area of 5750–6000 m, the decreasing rate is largest ($-0.58\%/year$). This decreasing trend is consistent with common speculation to the response of snow cover to global warming over the TP (Qin et al. 2006).

With only 7 years of available MODIS data, however, it is hard to arrive at any definitive conclusion regarding the linear trend of the snow-cover fraction for this area. A longer time series of data is needed to attain more robust conclusions.

9.4 Discussion and Conclusion

Based on the accurate and consistent MODIS snow cover fraction, snow cover distribution and variability over the TP is quantitatively explored for the first time. The snow cover distribution over the TP is extremely inhomogeneous due to complex terrain. The heavy snow cover concentrates on the western edge of the TP and Yarlung Zangbo Valley. In the interior of the TP, the snow cover fraction is relatively small but shows larger interannual variability. The highest snow cover fractions, typically in the range between 49 and 76%, are mostly concentrated at elevations higher than 6000 m. During the summer months (e.g., July and August), the Tibetan Plateau retains approximately 5% snow cover made up of scattered patches of snow.

The snow cover distribution linkage with terrain characteristics is also explored. In general, the SCF increases with elevation except in the interior of the TP. The SCF on southward and eastward slopes is commonly larger than on the northward and westward slopes. In the southward and eastward slopes, the SCF increases with the growth of steepness, but is not sensitive to incline in the northward and westward slopes. Snow cover in the concave and convex regions is relatively larger than in the flat regions. Annual evolution of snowlines shows distinguished characteristics in the different aspects. Special characteristics in the southward slopes are the results of specific configurations of available moisture and huge terrains. The snow cover ablation over TP begins in February at the lower elevation regions and delays with increasing of elevation. The decrement in the early spring is concentrated in the area of 2000–3000 m and moved to the area of 3000–5000 m after Apr and finally to the area above 4000 m. Due to large interannual variability, snow cover over the TP shows large anomalies or deviations from the climatological mean. These anomalies, however, are temporally limited, usually less than 3 months. The anomalies of snow cover do not persist into the summer.

There are slight decreasing trends in snow cover over TP during 2000–2006, especially in the maximum SCF during winter and in the higher elevation regions during the snow minimum season (August). A longer time series of data, however, needs to be examined to reach a definitive conclusion about temporal trends.

References

- Bamzai AS, Shukla J (1999) Relation between Eurasian snow cover, snow depth, and the Indian summer monsoon: an observational study. *J Climate* 12(10):3117–3132
- Che T et al (2003) Estimation of snow water equivalent from passive microwave remote sensing data (SSM/I) in Tibetan Plateau. In: *Proceedings-SPIE The International Society for Optical Engineering* (ed), International Society for Optical Engineering 1999, pp 405–412
- Qin DH et al (2006) Snow cover distribution, variability, and response to climate change in Western China. *J Climate* 19(9):1820–1833
- Frei A, Robinson DA (1998) Evaluation of snow extent and its variability in the Atmospheric Model Intercomparison Project. *J Geophys Res* 103(D8):8859–8872
- Hall DK et al (1999) Algorithm theoretical basis document (ATBD) for the MODIS snow-, lake ice-, and sea ice-mapping algorithms, algorithm theoretical basis document, Version, 4
- Hall DK et al (2001) Algorithm theoretical basis document (ATBD) for the MODIS snow and sea ice-mapping algorithms, NASA GSFC, September
- Hall DK et al (2002), MODIS snow-cover products. *Remote Sens Environ* 83(1):181–194
- Moore GWK (2004) Mount Everest snow plume: a case study. *Geophys Res Lett* 31(22):1–5
- Pu Z et al (2007) MODIS/Terra observed seasonal variations of snow cover over the Tibetan Plateau. *Geophys Res Lett* 34(6):6706. doi:10.1029/2007GL029262
- Pu ZX, Xu L (2009) MODIS/Terra observed snow cover over the Tibet Plateau: distribution, variation and possible connection with the East Asian Summer Monsoon (EASM). *Theor Appl Climatol* 97:65–278. doi:10.1007/s00704-008-0074-9
- Riggs GA et al (2000) MODIS snow products users' guide. NASA GSFC
- Riggs GA, Hall DK (2002) Reduction of cloud obscuration in the MODIS snow data product. In: *Proceedings of the 59 Eastern Snow Conference*, Stowe, Vermont USA, pp 4–6
- Thompson LG et al (1989) Holocene-late pleistocene climatic ice core records from Qinghai-Tibetan Plateau. *Science* 246(4929):474–477
- Walland DJ, Simmonds I (1996) Sub-grid-scale topography and the simulation of Northern Hemisphere snow cover. *Int J Climate* 16(9):961–982
- Wu GX et al (1995) The impacts of the snow-melt timing of the Tibetan Plateau on the seasonal variation in Southern China. *Gansu Meteor* 13:1–8
- Xu L (2007) Seasonal variations of snow cover over the Tibetan Plateau and ultimate effects on the East Asian summer monsoon. Thesis, University of Utah, 106 pp
- Yong L (1995) Effects of winter and spring snow cover over Tibetan Plateau on summer atmospheric general circulation over East Asia. *Plateau Meteor* 14(4):505–512

Chapter 10

The Global Geodetic Observing System (GGOS): Detecting the Fingerprints of Global Change in Geodetic Quantities

Hans-Peter Plag, Chris Rizos, Markus Rothacher, and Ruth Neilan

Abstract Modern geodetic observations from a wide range of space and terrestrial technologies contribute to our knowledge of the solid Earth, atmosphere, ocean, cryosphere, and land water storage. These geodetic observations record the “fingerprints” of global change processes and thus are a crucial independent source of high accuracy information for many global change studies. Many of the geodetic techniques require a globally distributed ground infrastructure, and associated space segment elements. In the past decade and half a variety of technique-specific services have been established under the auspices of the International Association of Geodesy (IAG) to facilitate global coordination of geodetic activities and to ensure the generation of high accuracy and reliable geodetic products to support geoscientific research. The Global Geodetic Observing System (GGOS) is an important component of the IAG, and is intended to be an “umbrella” for the IAG Services, with a primary coordinating function to ensure the development of an adequate global geodetic infrastructure, and a suite of integrated multi-technique products, that will meet the needs of scientific users. Coordination means bringing together the different geodetic observing techniques, services and analysis methods so as to ensure that the same standards, conventions, models and parameters are used in the data analysis and modelling of “Earth system” processes. Integration implies the combination of geometric, gravimetric, and rotational observations in data analysis and data assimilation, and the joint estimation and/or modelling of all the necessary parameters representing the difference components of the Earth system. The geodetic observations collected during the last decades have facilitated major scientific discoveries related to geohazards, climate and the global water cycle. Geodesy has the potential to contribute even more to global change studies, particularly if coordination and integration of the geodetic activities are continued.

H.-P. Plag (✉)
Nevada Bureau of Mines and Geology and Seismological Laboratory,
University of Nevada at Reno, Mail Stop 175, Reno, NV 89557, USA
e-mail: hpplag@unr.edu

10.1 Introduction

Humanity is increasingly being confronted with the limitations of a restless planet, with finite resources that cannot meet growing demands and a limited capacity to accommodate the impact of the increasingly dominant anthropogenic factor. The anthroposphere has grown into one of the most powerful factors in the Earth system, which is transforming the Earth's surface layers rapidly (Turner et al. 1990), and which is capable of changing major system processes. To illustrate the scale of the impact, consider the fact that more than 50% of the ice-free surface of the solid Earth has been transformed by humans. Among experts there is a broad consensus that humanity is changing the planet's climate (Doran and Kendall 2009). One could say that Earth has entered the geological epoch of the Anthropocene.

The Earth system is subject to a variety of dynamic processes driven by interior and exterior forces, operating on a wide range of temporal and spatial scales. As a consequence, large areas of the Earth's surface are exposed to natural hazards, including geohazards, storms, storm surges, and floods. Urban settlements are sprawling into areas at high risk from such natural hazards, thus increasing the vulnerability of society. At the same time, natural and anthropogenic climate change, is changing the hazards and introducing a large uncertainty into any predictions of future events and developments. Considering the scale of the anthropogenic impact on the Earth system, it is clear that decisions made today will influence the well-being of future generations. Yet our limited capability of predicting the consequences of our activities, including effectiveness of different mitigation strategies, hampers the development of mitigation and adaptation policies.

A number of recent World Summits have emphasized the paramount importance of sustainable development for a prosperous future of the anthroposphere and confirmed the necessity to reach to an operational implementation of the ethical concept of Sustainable Development as proposed by the World Commission on Environment and Development (1987). Understanding of the major processes in the Earth system, and its changes over time, is one of the many prerequisites required for this operationalization, which depends on both the capability to monitor the state of the Earth system, and the capability to predict the future trajectories of the system under various assumptions. Both capabilities cannot be developed without a comprehensive monitoring of the Earth system. Recognizing the urgent need for a coordinated and sustained program of Earth observation, the recent Earth Observation Summits (EOS) have tasked the intergovernmental Group on Earth Observations (GEO) with the implementation of the Global Earth Observation System of Systems (GEOSS). GEOSS is designed with a focus on nine Societal Benefit Areas (SBAs) (GEO 2005), including the SBAs of Climate, Weather, and Disasters.

Geodesy provides the foundation on which most Earth observation systems are built. In this function, the geodetic observing system is essential for Earth observation in general and GEOSS in particular. Geodesy is the science of measuring and mapping the geometry, rotation and gravity field of the Earth including their variations with time. These three characteristics of the Earth system are inherently

related to the dynamics of the system, and to mass and energy transport throughout the system.

The development of space-geodetic technologies has revolutionized the methods of geodesy, and make possible the comprehensive monitoring of physical characteristics of the Earth system with high accuracy and resolution. Since the advent of space-geodesy about half a century ago, the accuracy of positioning in a global geodetic reference frame has increased by roughly an order of magnitude every decade, reaching today sub-centimetre for relative positions on global scales and sub-millimetre per year changes in these positions. With today's space-geodetic techniques, changes in the surface of the solid Earth, the oceans, land surface waters, and the ice sheets, can be measured with unprecedented accuracy and with ever increasing spatial and temporal resolution. These geodetic observations provide critical information on geodynamic processes underlying geohazards such as earthquakes, volcano eruptions, landslides, and subsidence, as well as changes in the global water cycle such as melting of ice sheets, sea level rise, and changes in land water storage. Variations in Earth rotation are inherently related to the global dynamics of the coupled atmosphere-ocean-solid Earth system. For observations of the time-variable Earth rotation, accuracy has also increased several orders of magnitude over the last few decades. These observations are not only critical for our understanding of the processes in the core and mantle of the solid Earth but also provide important constraints on large-scale climate phenomena and climate models. Dedicated gravity satellite missions have improved our quantitative knowledge of Earth's static gravity field to a level where they now provide vital constraints for ocean circulation models (Shum et al. 2010). Satellite observations of the time-variable gravity field, for the first time, provide estimates of the changes in water storage on subcontinental scales and with temporal resolutions reaching 10 days (see Plag et al. (2009a) for references). The combination of the geodetic observations has allowed the determination of a global terrestrial reference frame with centimetre accuracy, an internal precision at the sub-centimetre level, and a long-term stability on the order of 1 mm/yr or better.

Most of the technological developments in space-geodesy, and the rapidly improved scientific understanding of the observations, have been facilitated by the scientific expertise of the global geodetic community gathered in the International Association of Geodesy (IAG). Over the last 15 years, the coordination of the global geodetic infrastructure, data collection and processing, and product development has come to a large extent through a number of technique- and product-specific IAG Services. These today provide valuable observations and products not only to the science community but also for a wide range of non-scientific applications. The International GNSS Service (IGS), established in 1994, provided the prototype model for these services. Recognizing the need to have a common voice for the increasing number of IAG Services as well as a link between IAG as a whole and relevant international Earth observation and research programs, the IAG initiated the Global Geodetic Observing System (GGOS) during the IUGG meeting in 2003 as an IAG Project. After an initial definition phase and the implementation of core

elements of the organizational structure of GGOS, at the IUGG meeting in 2007 the IAG elevated GGOS to the status of permanent observing system of IAG.

The acronym “GGOS” has two very distinct aspects, which should be clearly distinguished: (1) the “organization GGOS”, which consists of committees, panels, bureaus, and working groups, and (2) the “observation system GGOS”, which comprises the infrastructure of a wide range of instrument types, satellite missions, and data and analysis centres (Plag et al. 2009b). While GGOS as an organization has established its structure from essentially new entities, the infrastructure for GGOS is provided to a large extent by the existing IAG Services. The organizational components of GGOS are an integral part of the IAG structure (see, for example, Beutler et al. (2009) for a description of GGOS as an organization). With respect to the physical observing system [see, for example, Rothacher et al. (2009) for a description of GGOS as a system], GGOS is facing a rather different situation: a considerable part of the infrastructure is provided by organizations that are not formally (or only loosely) affiliated with the IAG. The GPS system, for example, which is crucial for the IGS, is owned and maintained by the U.S. Department of Defense. Satellite missions contributing to GGOS, including satellite altimetry, satellite gravimetry, and InSAR, are implemented and operated by space agencies, while GGOS only utilizes the observations. Dedicated infrastructure such as the stations in the global tracking networks, the analysis centres, and even the components of the service organizations are provided on a voluntary basis. Consequently GGOS as an observing system depends crucially on infrastructure provided by others, and GGOS as an organization faces the challenge of ensuring, through dialogue with the relevant providers, that this infrastructure is available on a continuous basis. Considering the fact that the geodetic reference frames, which are made available through the IAG Services and GGOS, provide the backbone for geospatial services that underpin many modern societal applications, GGOS thus faces the challenge of promoting a core, but often invisible, element of the infrastructure of a modern society to those who, on the one hand, can provide the resources to sustain this infrastructure, and, on the other hand, can benefit from the geodetic observations and products.

10.2 Geodesy’s Contribution to Earth Observation

Global geodesy serves Earth observation and global change science in two distinct ways: (1) geodesy provides the reference frames required for all location-dependent observations and thus contributes to the foundation of most Earth observations, and (2) geodesy provides observations of the time-variable shape, gravity field, and rotation of the Earth, and thus contributes to the Earth observation database. Within geodesy, the observations and the reference frames are inherently coupled. In fact, since the solid Earth is constantly deformed by internal geodynamic processes and by mass and energy transport in the fluid envelope of the solid Earth, the capability to understand and model these processes ultimately determines the accuracy of the global geodetic reference frames.

10.2.1 The Global Geodetic Reference Frames

Assigning time-dependent coordinates to points and objects, and describing the motion of the Earth in space requires access to an appropriate reference system. In geodesy, two global reference systems are basic for this purpose, namely the celestial reference system and the terrestrial reference system, which are dynamically linked to each other by the Earth's rotational motion [see, e.g., Plag et al. (2009a) for references]. The two most accurate reference systems currently available are the International Celestial Reference System (ICRS) and the International Terrestrial Reference System (ITRS), which are conventional coordinate systems including all conventions for origin and orientation of the axes, scale, and physical constants, models, and processes to be used in their realization. These two systems are realized through their corresponding "reference frames," i.e., the International Celestial Reference Frame (ICRF; a set of estimated positions of extragalactic reference radio sources) and the International Terrestrial Reference Frame (ITRF; a set of estimated positions and velocities of globally distributed reference marks on the solid Earth's surface), respectively. These two frames are linked to each other by estimates of the Earth Orientation Parameters (EOPs). Based on the observations and analysis results provided by the IAG Services, the International Earth Rotation and Reference Systems Service (IERS) determines and provides access to the ITRF, the ICRF, and the EOPs.

The ITRF is today the most accurate realization of a global geodetic reference system. This well-defined, long-term stable, highly accurate and easily accessible reference frame is the basis for all precise positioning on and near the Earth's surface. The ITRF underpins all geo-referenced data used by society for so many uses, and it is the basis for most other reference frames, including WGS84 and the reference frame GTRF for the GALILEO system. Because of this, the ITRS and the ITRF are not only indispensable for Earth observation in general, but also for many societal applications ranging from navigation, mapping, surveying, national and regional reference frames, to civil engineering.

10.2.2 Role of Geodetic Observations for Science

As mentioned above, geodesy is the science of determining the geometry, gravity field, and rotation of the Earth, and their evolution in time. Modern geodetic techniques permit the measurement of changes in the geometry of the Earth's surface with an accuracy of millimetres over distances of several 1000 km. The secular surface kinematics can be monitored with sub-mm/yr accuracy. Geodetic imaging techniques increasingly gain importance, particularly when integrated with the traditional point-based approach of geodesy. Based on these geometric techniques, temporal changes in the Earth's shape, rotation and gravity field are provided with increasing spatial and temporal resolution, increasing accuracy, and with decreasing latency. The internationally coordinated global geodetic station networks provide a continuous monitoring of the ITRF and the EOPs. In combination with ITRF, the

Global Navigation Satellite Systems (GNSS) provide access to precise point coordinates in ITRF anytime and anywhere on the Earth's surface with centimetre-level accuracy.

Data from a large number of space missions and other geodetic observations have contributed to the determination of models of the Earth's gravity field. The significant improvement in the resolution and precision of our knowledge of the Earth's gravity field is due to the satellite missions CHAMP and GRACE, in orbit since 2000 and 2002, respectively. In particular GRACE has pushed our knowledge of the static geoid to centimetre-level accuracy. The European GOCE mission, launched on 17 March 2009, will further improve precision and resolution of the static part of the gravity field to an unprecedented level. The GRACE mission is designed to monitor changes in Earth's gravity field induced by mass changes at spatial scales less than 500 km and with amplitudes of the order of 1 cm of water equivalent.

All these geodetic observations record the "fingerprints" of mass movements in the solid Earth, ocean, atmosphere, ice sheets and terrestrial water storage; they allow the determination of the displacement field associated with earthquakes and the velocity and strain fields of the Earth's surface; they scale mass and volume changes in the ocean; they monitor the changes in ice sheets; they sense the changes in land water storage; they measure the water content in the atmosphere; and they provide crucial constraints for all models of geophysical processes in the Earth system [see Plag et al. (2009a) for references].

At the accuracy level reached today in space-geodesy, the variations in the Earth's shape, gravity field, and rotation are caused by a number of Earth system processes acting on a wide range of spatial and temporal scales. Meaningful interpretation of the geodetic observations therefore requires the development of models that can predict the geodetic observations for an increasing number of these processes with the same accuracy as, or better than, the observations. In the course of this process, geodesy has developed into a science making unique contributions to the understanding of the Earth system, its dynamics, and its response to climate change. Consequently, with modern instrumentation and analytical techniques, the scope of geodesy is rapidly extending from a mere monitoring of the geometric, gravitational and rotational changes, to the modelling and understanding of the causes of the observed changes. With this broader scope, geodesy contributes not only to the scientific understanding of the Earth system but also the functioning, and security of society in general.

10.3 GGOS: A Multi-technique, Multi-layered Yet Integrated System

10.3.1 A Value-Chain from Observations to Applications

GGOS provides products that are pivotal for Earth observation, Earth science, geo-information systems, and terrestrial and planetary navigation. From a value-chain point of view, connecting observations at one end of the chain to user applications

at the other end, GGOS comprises four main components. (1) The instrumentation includes global terrestrial networks of geodetic stations and observatories, Earth observing satellites, satellite navigation systems, and planetary missions. (2) The data infrastructure comprises the infrastructure for data transfer, data management and archiving, and data and product dissemination. (3) The data analysis covers the complete and consistent data processing chain from the initial acquisition and the processing of large amounts of observations, to the consistent integrated analysis and combination of products, and the assimilation of the observations into complex models of the Earth system or components of this system. (4) The GGOS Portal provides a unique access point for GGOS users to all GGOS products, including relevant metadata and documentation.

GGOS faces the challenge to develop the geodetic technologies and the observing system so that they meet the demanding and evolving observational requirements in terms of reference frame accuracy and availability, as well as in terms of spatial and temporal resolution and accuracy of the geodetic observations. GGOS has assessed these requirements for science, Earth observations and societal application (see Gross et al. 2009). The most demanding observational requirements result mostly from scientific applications. For reference frame accuracy, scientific studies of sea level change caused by climate change appear to pose the most demanding requirements of accuracy and stability at the levels of 1 mm and 0.1 mm/yr, respectively. For the geoid, for use in ocean general circulation models to define the mean sea surface topography, and the use with GNSS for the determination of heights at the millimetre level requires the static geoid to be accurate and stable at levels of 1 mm and 0.1 mm/yr, respectively. For Earth orientation parameters, the most demanding application is likely to be the tracking and navigation of interplanetary spacecraft, which is a capability-driven application requiring the most accurate EOPs that can be determined, such that they are consistent in accuracy with the accuracy of the ITRF and ICRF.

A key element in this value chain is the theory used for the modelling and interpretation of the geodetic observations. Much of the theory used today was developed at a time when geodetic observations were at a much lower accuracy level, and parts of the theory is not accurate enough to fully exploit the current, and much less so future improved, accuracy. An example are the satellite gravity observations, where theory development lags behind observation accuracy (Xu 2008), although efforts are made to bring theory in line with the improved accuracy (e.g., Mayer-Gürr et al. 2005). Users of geodetic products will have to consider carefully any mismatch in theory and observations, and GGOS will have to focus on bringing theory in line with a constantly improving accuracy of the observations.

10.3.2 A System-of-Systems

From a system-oriented point of view, the global geodetic infrastructure can be characterized as a multi-technique “system-of-systems”, which observes key geometric and physical quantities of the Earth system with specialized technologies

Table 10.1 The global geodetic observing system (GGOS)

Component	Objective	Techniques	Responsibility
I. Geokinematics (size, shape, kinematics, deformation)	Shape and temporal variations of land/ice/ocean surface (plates, intra-plates, volcanos, earthquakes, glaciers, ocean variability, sea level)	Altimetry, InSAR, GNSS-cluster, VLBI, SLR, DORIS, imaging techniques, levelling, tide gauges	International and national projects, space missions, IGS, IAS, future InSAR service
II. Earth Rotation (nutation, precession, polar motion, variations in length-of-day)	Integrated effect of changes in angular momentum and moment of inertia tensor (mass changes in atmosphere, cryosphere, oceans, solid Earth, core/mantle; momentum exchange between Earth system components)	Classical astronomy, VLBI, LLR, SLR, GNSS, DORIS, under development: terrestrial gyroscopes	International geodetic and astronomical community (IERS, IGS, IVS, ILRS, IDS)
III. Gravity field	Geoid, Earth's static gravitational potential, temporal variations induced by solid Earth processes and mass transport in the global water cycle	Terrestrial gravimetry (absolute and relative), airborne gravimetry, satellite orbits, dedicated satellite missions (CHAMP, GRACE, GOCE)	International geophysical and geodetic community (GGP, IGFS, IGeS, BGI)
IV. Terrestrial Frame	Global cluster of fiducial points, determined at mm to cm level	VLBI, GNSS, SLR, LLR, DORIS, time keeping/transfer, absolute gravimetry, gravity recording	International geodetic community (IERS with support of IVS, ILRS, IGS, and IDS)

VLBI Very Long Baseline Interferometry, *SLR* Satellite Laser Ranging, *LLR* Lunar Laser Ranging, *GNSS* Global Navigation Satellite Systems, *DORIS* Doppler Orbitography and Radio Positioning Integrated by Satellite, *InSAR* Interferometric Synthetic Aperture Radar, *IGS* International GNSS Service, *IAS* International Altimetry Service, *IVS* International VLBI Service for Geodesy and Astrometry, *ILRS* International Laser Ranging Service, *IDS* International DORIS Service, *IERS* International Earth Rotation and Reference Systems Service, *IGFS* International Gravity Field Service, *GGP* Global Geodynamics Project, *BGI* International Gravimetric Bureau, *IGeS* International Geoid Service.

Modified from Plag et al. (2009a)

(Table 10.1). The technique-specific IAG Services take care of the coordination of infrastructure relevant for a specific technique Rothacher et al. (2009). Coordination between techniques and combination of products from different techniques to higher-level products is undertaken by the IERS and the International Gravity Field Service (IGFS).

10.3.3 A Multi-layered System

Looking from the Earth’s surface upward, the global geodetic infrastructure appears to be layered and can be characterized by five major levels of instrumentation and objects that actively perform observations, are passively observed, or both (Fig. 10.1). These levels are (Rothacher et al. 2009):

- Level 1: the terrestrial geodetic infrastructure, which comprises ground stations of in situ and tracking networks, observatories, communication links, and infrastructure for data storage, processing, modelling, and distribution;
- Level 2: the Low-Earth Orbiting (LEO) satellite missions, which include different types of altimeters, synthetic aperture radar, gravity satellite missions, Light Detection And Ranging (LIDAR) sensors, and other instruments;
- Level 3: the satellites of the Global Navigation Satellite Systems (GNSS) and the LAGEOS-type Satellite Laser Ranging (SLR) satellites;
- Level 4: the planetary missions and geodetic infrastructure on Moon and planets, including reflectors for Lunar Laser Ranging (LLR);
- Level 5: the extragalactic radio-signal emitting quasars.

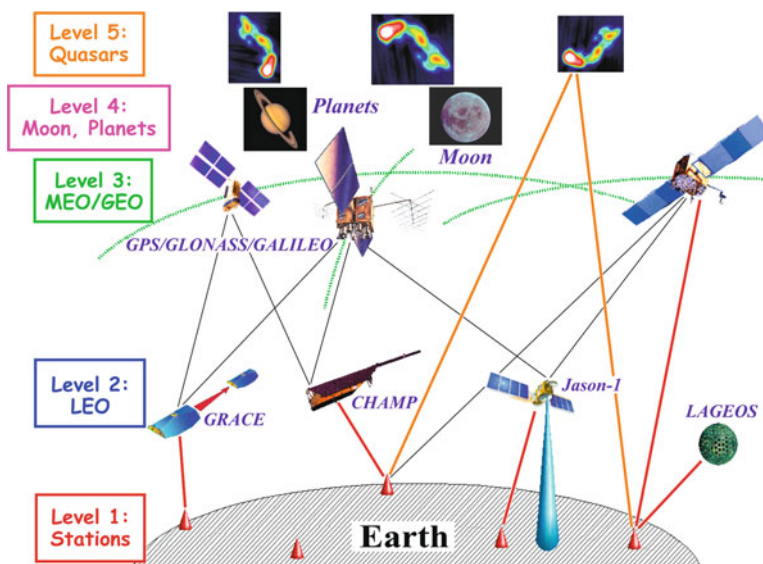


Fig. 10.1 The five layers of the global geodetic infrastructure. From Rothacher et al. (2009)

10.3.4 Integration Through Multiple Links

Independent of whether the instruments and objects in a layer are active or passive, receivers or emitters or both, these five levels of instrumentation and objects

are connected in various ways. Firstly, all geodetic techniques measure the “output” of the same unique Earth system and are thus affected by the same environmental processes, although with different sensitivities and noise sources. Most, if not all, physical processes in the Earth system are associated with mass redistributions (see Section 10.4). The geodetic fingerprints of these mass redistributions and changes in the system’s dynamics are captured by the different techniques according to their specific sensitivity. For example, displacements of the surface of the solid Earth, oceans, ice sheets, and land water areas are observed with several independent geometric and imaging techniques. Earth rotation variations are derived from several independent techniques and through combinations of these techniques; and changes in surface mass storage are extracted from observations of the time-variable gravity field, Earth rotation, and surface displacements. Thus, the different parts of the overall system are cross-linked through observations of the same geometric and physical quantities, and, to a certain extent, they are inter-dependent. In the data processing, the geodetic observations are further linked through geophysical models of the system processes. Therefore, consistency of data processing, modelling, and conventions across the techniques and across the three main areas of geodesy is mandatory if the full potential of the observing system is to be realized.

Most importantly, for the integration of the individual layers and systems into a consistent observing system, the different techniques are connected through co-location of sensors both on Earth and in space. For the separation of technique-dependent effects from the fingerprints of the Earth system processes in the geodetic observations, co-location of different and often complementary techniques is crucial. Co-location of different techniques at the same location on Earth has vital for ensuring the stability of the ITRF. Today, the uneven geographical distribution of “core” stations with three or more space-geodetic techniques is a limiting factor for the accuracy of the ITRF. GGOS must work towards a more even distribution of core stations.

The emerging co-location of different sensors and observation types onboard a satellite is also extremely important to establish connections between the observation techniques (Rothacher et al. 2009). This in-space co-location complements the co-location of stations on the Earth. An example is the rapid progress achieved in orbit determination with the tracking data of the TOPEX/Poseidon satellite using DORIS, GPS, SLR and altimetry crossovers. GGOS is therefore promoting that future satellite missions put emphasis on establishing links between different observation and tracking techniques. For example, it is of particular importance that all GNSS satellites be equipped with laser retro-reflectors arrays, thus providing a link between GNSS and SLR (Plag et al. 2009a; Rothacher et al. 2009).

10.3.5 An Integrated System Sensing Atmosphere, Hydrosphere, and Solid Earth

Facilitated by these links through physics, models and co-location, the five layers of the global geodetic infrastructure form the complex, integrated GGOS as

an observing system. The major observation types acquired by this system include: (1) observations of the microwaves emitted by GNSS satellites at the ground stations and at LEO satellites; (2) laser ranging to LEOs, dedicated laser ranging satellites, GNSS satellites and the Moon; (3) observation of the microwaves emitted by quasars using antennas in the Very Long Baseline Interferometry (VLBI) network; (4) instrumentation onboard the LEO satellites measuring accelerations, gravity gradients, satellite orientation, etc.; (5) radar and optical observations of the Earth's surface (land, ice, glaciers, sea level, ect.) from remote sensing satellites; (6) distance measurements between satellites (K-band, optical, interferometry, etc.); (7) measurements of other quantities on the ground related to variations in gravity (with absolute and relative gravimeters), Earth rotation (for example, with ring lasers), sea level (for example, with tide gauges), and water storage (for example, lake groundwater levels) (see Rothacher et al. (2009) for more details). For many of these observations, in particular, the space-geodetic observations, time measurements are crucial and, ultimately, the accuracy of the geodetic observations depends on the accuracy of time measurements.

With these observations, the geodetic infrastructure not only measures time-variable Earth shape, gravity field, and rotation. As they propagate through the ionosphere and troposphere, the microwaves are refracted by the atmosphere and thus sense this component of the Earth system as well.

10.4 Global Change Results

Many of the burning questions related to the water cycle, climate variability, global change, and geohazards cannot be solved without knowledge of energy and mass transports throughout “System Earth” and the associated dynamics (Rummel et al. 2009). Most of the processes leading to mass and energy transport affect the Earth's figure (geometry), its gravity field and its rotation. Consequently, geodetic observations record the fingerprints of global change processes. However, the signals induced by global change in geodetic observations are generally small (of the order of parts-per-billion of the quantities), and they are often embedded in larger variations not caused by global change. Therefore, in addition to geodetic observations with an accuracy considerably better than these signals, identifying and quantifying the global change signals also requires accurate modelling of all known processes in an Earth system model (Herring et al. 2009; Rummel et al. 2009).

Many scientific studies of global processes, including those related to global and climate change, have already benefited from detailed knowledge of the Earth's time-variable shape, rotation, and gravity field. Over the last fifteen or so years, the global geodetic networks have provided an increasingly detailed picture of the temporal variations in the Earth's shape and of the kinematics of points on the Earth's surface, including the ocean, ice cover, and land surfaces. The observations have been used, for example, to determine improved models of the secular velocity field as an input for studies of plate tectonics, post-glacial rebound, sea level changes, and ice load changes, to derive models of the global strain rate field (Fig. 10.2), to study

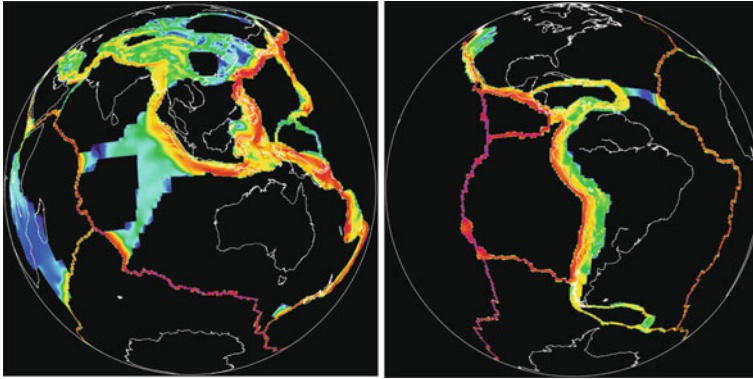


Fig. 10.2 Geodetic observations of Earth's time-variable shape are used, for example, to determine strain rate models. Shown is the second in-invariant of a strain rate tensor model obtained from global GPS observations. C. Kreemer 2009, personal communication

seasonal loading and derive seasonal variations in the terrestrial hydrosphere, to invert for mass changes in ocean, ice sheets and land water storage, to improve the modelling of the seasonal term in polar motion, and to study transient surface deformations prior to, during and after earthquakes. Geodetic techniques provide the means to observe surface deformations on volcanoes, in unstable landslide prone areas, associated with earthquakes and fault motion, or subsidence caused by anthropogenic activities such as groundwater and oil extraction (for more examples and references, see Plag et al. 2009a). In the near future, geodetic observing techniques will be able to determine the magnitude and displacement field of great earthquakes in near real-time as support for early warning systems (e.g., Blewitt et al. 2006).

Variations in Earth's rotation are induced by mass transport in the Earth system and the exchange of angular momentum among its components. Earth rotation observations have provided insight into many global-scale dynamic phenomena, including sea level changes, post-glacial rebound, and hemispheric seasonal changes in land water storage and snow load. An example at interannual time scales is the El Niño/Southern Oscillation (ENSO) phenomenon, which is the most prominent feature of the climate system at these time scales. ENSO events are associated with a collapse of the tropical easterlies leading to an increase of the atmospheric angular momentum, which in turn is compensated for by a decrease of the solid Earth's angular momentum and an increase of the length-of-day of up to 0.5 milliseconds (for particularly strong ENSO events). ENSO events are associated with significant and far-reaching impacts on interannual climate, and early detection and forecasting of the timing and magnitude of these events is of high societal relevance. Geodetic observations of Earth rotation and sea surface height anomalies provide pivotal constraints for both detection and forecasting (for references, see Rummel et al. 2009).

Gravity field variations are caused by many different processes in the Earth system (Fig. 10.3), and most of these processes have geodetic fingerprints above the

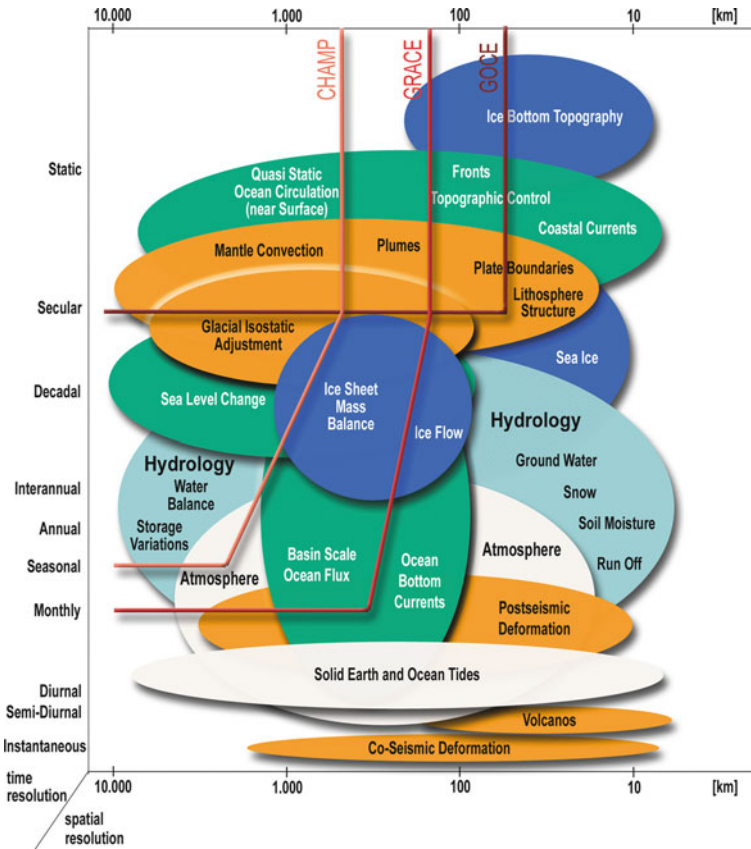
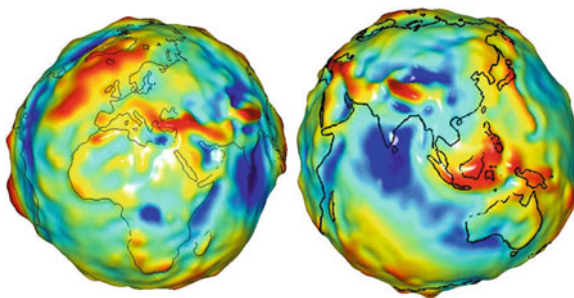


Fig. 10.3 Mass transport in the Earth system. The different processes are associated with characteristic spatial scales, ranging from a few km to global, and temporal scales from instantaneous to secular and static. The capability of current dedicated gravity satellite missions to capture the “fingerprint” of some of these processes is indicated. From Ilk et al. 2005

accuracy level reached by space-geodesy. For example, gravity effects of co-seismic and post-seismic deformation associated with large earthquakes have been identified in GRACE observations (e.g., Bao et al. 2005; Panet et al. 2008). At time scales from weeks to several decades, the largest mass redistributions on the surface of the solid Earth occur in the hydrological cycle, and the GRACE mission is providing unprecedented insight into the water cycle down to spatial scales of about 500 km and on sub-monthly time scales (Fig. 10.4; for references, see Plag et al. 2009a).

Figure 10.3 illustrates that the different processes are associated with different temporal and spatial scales, which lead to multiple geodetic signatures. Mass movements in the hydrosphere have temporal scales from weeks to decades, and spatial scales from 10 km to continental or basin scales. These mass movements load and deform the solid Earth, impact on the gravity field of the Earth system, and change

Fig. 10.4 Geodetic observations of the time-variable gravity field of the Earth. The pictures show a snapshot of monthly changes in the Earth's gravity field with respect to the mean field. The changes are caused mainly by mass transport in the fluid envelope of the solid Earth. Courtesy NASA



the rotation of the solid Earth. The response of the solid Earth to the loads includes an elastic response to concurrent mass redistributions, as well as a viscous response to past mass movements. The response in Earth's shape, gravity field and rotation depends on the spatial scale of the loads, with surface displacements being more sensitive to smaller loads than gravity and rotation. Consequently, the different geodetic techniques have specific temporal and spatial sensitivities. Sampling of the complete spatio-temporal spectrum with high resolution and accuracy with a set of different techniques is therefore necessary in order to be able to invert the geodetic observations to derive mass changes associated with the individual Earth system processes. For example, inversion of surface displacements observed with GPS and geophysical models of these displacements have errors at the level of a few centimetres at a few hundred km wavelength, while mass estimates derived from GRACE at 200 km have errors larger than 20 cm. A combination of integrated geodetic observations with models of the water cycle is therefore a promising approach. For example, results of combined inversion of observations from GRACE, SLR, VLBI, GPS and predictions of an ocean model (ECCO) show a dramatic melting of Greenland (Wu 2008, personal communication).

There are many climate-related scientific problems that can be addressed using combined geodetic observations. For example, one of the projects in the Global Energy and Water Cycle Experiment (GEWEX) addresses the question of how changes in land use and aerosol may have impacted monsoons (Lawford 2008, personal communication). Geodetic observations can provide important constraints on the associated changes in land water storage. Other areas are extremes: there are periods of "quiet" variations and then there are periods of extreme variations, and geodetic observations may help to characterize and potentially predict these periods. For example, the forecasts of the drought conditions in 2005 in Canada failed completely (Lawford 2008, personal communication), but using global geodetic observations to characterize the state of the Earth system may help to improve such forecasts. There is also an urgent need to improve season-by-season prediction. On the one hand, geodetic observation support the validation of numerical (weather and climate) prediction models, and, on the other hand, assimilation of these observations into prediction models may in the future improve the predictive capabilities of these models.

Combination of the full suite of geodetic observations with land surface hydrological models helps to increase spatial resolution. An important advantage of the geodetic observations is that they are not limited in terms of penetration depth: subsurface mass changes have their fingerprint in all geodetic observations. Current challenges for an operational use of the geodetic observations, for example, in water management, drought and seasonal predictions, include considerable differences between observed and modelled variables, a spatial resolution of observations that is too coarse, a temporal resolution that is too low, and a latency that is too high.

The gravity field determination from dedicated space missions has also contributed tremendously to advances in a number of fields of geodesy, including reference frames, the orbits of ocean radar altimetry satellites, laser altimeters, and sea surface topography from satellite altimetry [see Plag et al. (2009a) for references]. The integration of various satellite missions with the geometric techniques created new opportunities for the study of mass transport in the Earth system in a globally consistent way (Ilk et al. 2005). Oceanographic applications illustrate the unique way in which the combined geodetic observations provide accurate and quantitative constraints on the ocean mass budget, tidal dissipation, near-surface ocean flow and its variability, and large-scale ocean mass variations (Cazenave et al. 2010; Shum et al. 2010). The observations are invaluable for understanding the causes of sea level rise and the dynamics of ocean mass redistribution. Geodetic observations are also crucial in reducing the current large uncertainties in future global and local sea level predictions, a key issue hampering the development of urgently needed adaptation policies for one of the most severe impacts of global warming: coastal inundation (Cazenave et al. 2010; Plag et al. 2010).

In order to fully exploit the complimentary nature of the changes in Earth's shape, gravity field and rotation for global change studies, it will be necessary to develop the observing system based on our understanding of the geodetic fingerprints of these changes. However, currently other factors have a greater impact on the development of geodetic infrastructure. As an example, we compare areas of large on-going mass changes to the network of GPS stations. The secular trends in surface mass as determined from GRACE is shown in Fig. 10.5. These mass changes induce both gravity changes and surface displacements, with the ratio of surface displacements to gravity changes increasing with decreasing spatial scales. Therefore, observations of surface displacements, particularly for areas where GRACE senses large secular mass variations, would be complementary to the gravity observations. However, comparing the locations of GPS stations for which observations are available in public data archives, it is obvious that only some areas with large mass decreases are well covered with GPS stations (e.g., North America), while others are not (e.g., Greenland, Africa, South America, and Antarctica, see Fig. 10.5). Future deployments of GNSS sites should therefore aim to reduce the mismatch between areas with mass changes and GNSS coverage.

In the future, new measurement techniques are likely to evolve and these will be included into GGOS. An example is GNSS reflectometry, which is based on recording GNSS signals onboard LEO satellites after they are reflected off the ocean or ice surfaces. In 2013, more than 120 GNSS satellites will be providing signals for this

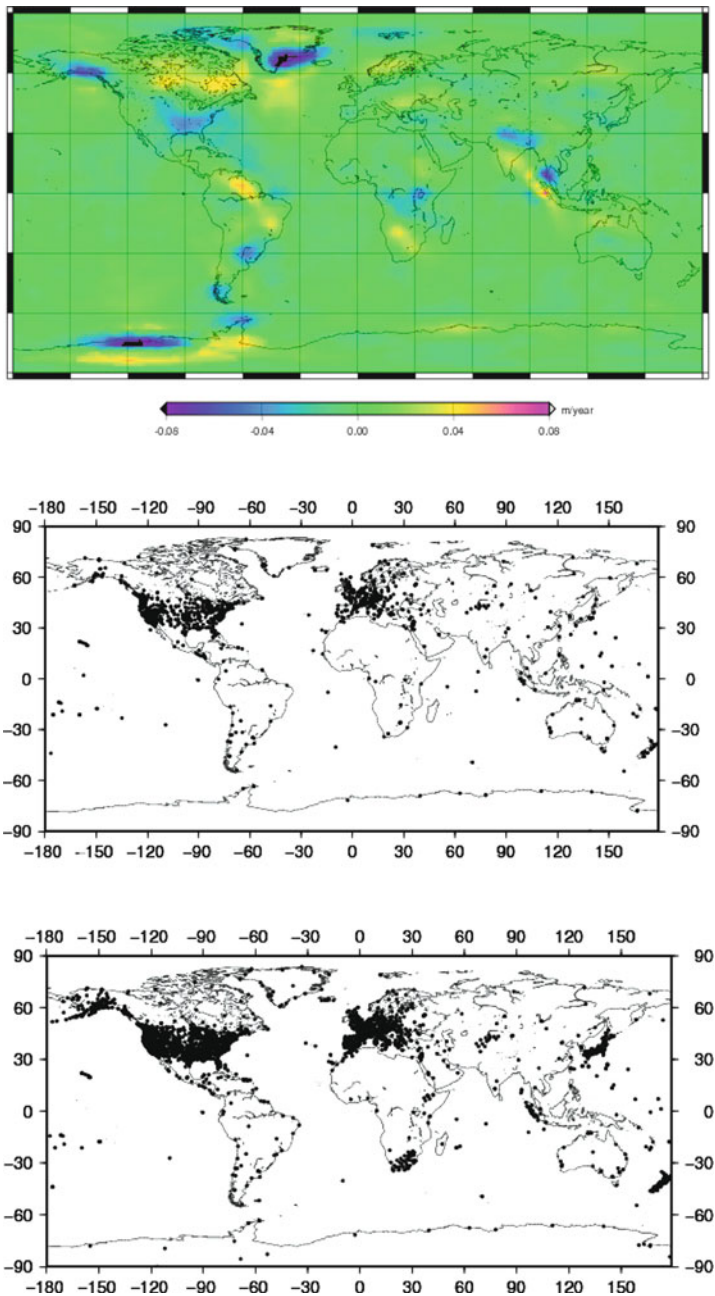


Fig. 10.5 *Top:* Mass changes in m/year derived from GRACE. From Watkins et al. (2008). *Middle and bottom:* Location of GPS sites the Nevada Geodetic Laboratory (NGL) has discovered in public data archives and for which NGL is downloading and analyzing data routinely. *Middle:* Location of 1299 GPS sites available in 2002. *Bottom:* Location of 3825 GPS sites available in 2008 (GPS site locations from Kreemer and Blewitt 2008, personal communication)

technique. Therefore this concept has a huge potential to measure ocean and ice surface topography and roughness with high spatial and temporal resolution, assisting in ocean and ice monitoring and for global tsunami early warning systems. Another emerging concept is the co-location of space geodetic techniques on special micro-satellites. Such micro-satellites could be equipped with GNSS antennas for precise orbit determination and for radio occultation studies, star sensors for attitude, an SLR retro-reflector and a VLBI signal transmitter. As the satellite (or satellite constellation) orbits the Earth it would “connect” all the global core sites and improve co-location from space.

10.5 Future Developments

In the future, new measurement techniques are likely to evolve and these will be included into GGOS. An example is GNSS reflectometry, which is based on recording GNSS signals onboard LEO satellites after they are reflected off the ocean or ice surfaces. In 2013, more than 120 GNSS satellites will be providing signals for this technique. Therefore this concept has a huge potential to measure ocean and ice surface topography and roughness with high spatial and temporal resolution, assisting in ocean and ice monitoring and for global tsunami early warning systems. Another emerging concept is the co-location of space geodetic techniques on special micro-satellites. Such micro-satellites could be equipped with GNSS antennas for precise orbit determination and for radio occultation studies, star sensors for attitude, an SLR retro-reflector and a VLBI signal transmitter. As the satellite (or satellite constellation) orbits the Earth it would “connect” all the global core sites and improve co-location from space.

Currently, temporal inhomogeneities of geodetic time series due to technological developments, changing station distribution, and improvements in geophysical models and geodetic analysis methods are a limitation for many global change studies. Developments in analysis capabilities through new approaches for single techniques (such as, for example, the recent ambizap algorithm, see Blewitt 2008) or combined analyses, improvements of the reference frame (for example, Gross et al. 2009) and the geophysical models, as well as increased computer storage and processing power will provide the necessary capabilities for rapid and homogeneous reprocessing of the complete global geodetic database with the highest possible accuracy in order to provide datasets for global change and climate studies that are as homogeneous in time as possible. A prerequisite for exploiting the full potential of geodesy for Earth observation, Earth system monitoring, and for many practical applications is, however, a sophisticated integration of all geodetic techniques (spaceborne, airborne, marine and terrestrial), processing models and geophysical background models into one system. This integration will permit – as part of global change research – the assessment of surface deformation processes and the quantification of mass anomalies and mass transport inside individual components, and mass exchange between the components of the Earth’s system. GGOS is pivotal in facilitating this integration.

Acknowledgments The authors are grateful to the IAG Community, the Services, Commissions, the GGOS Steering and Executive Committees, the Science Panel, and the Working groups: GGOS is built by the best effort of the many individuals in these IAG and GGOS components. GGOS also depends on the continuous support of many other contributors, in particular the space agencies, which provide infrastructure crucial to GGOS. The GPS site locations for the maps in Fig. 10.5 were provided by Corne Kreemer and Geoff Blewitt. Part of the work carried out by the lead author was supported by several NASA grants under the ROSES program and by a JPL contract.

References

- Bao LF, Piatanesi A, Lu Y, Hus HT, Zhou XH (2005) Sumatra tsunami affects observations by GRACE satellites, *Eos. Trans. Am. Geophys. Union* 86:353, 356
- Beutler B, Pearlman M, Plag H-P, Neilan R, Rothacher M, Rummel R (2009) Towards GGOS in 2020. In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, pp 273–282
- Blewitt G (2008) Fixed-point theorems of GPS carrier phase ambiguity resolution and their application to massive network processing: ‘Ambizap,’ *J. Geophys. Res.* 113:B12410. doi:10.1029/2008JB005736
- Blewitt G, Kreemer C, Hammond W, Plag H-P, Stein S, Okal E (2006) Rapid determination of earthquake magnitude using GPS for tsunami warning systems. *Geophys. Res. Lett.* 33:L11309. doi:10.1029/2006GL026145
- Cazenave A, Chambers DP, Cipollini P, Fu LL, Hurell JW, Merrifield M, Nerem RS, Plag H-P, Shum CK, Willis J (2010) Sea level rise: regional and global trends, Plenary Paper. In: *Proceedings of OceanObs’09, September 2009, Venice, Italy*. In press
- Doran PT, Kendall M (2009) Examining the scientific consensus on climate change, *Eos. Trans. Am. Geophys. Union* 90:22–23
- GEO (2005) Global earth observing system of systems GEOSS – 10-year implementation plan reference document – Draft}, Tech. Rep. GEO 1000R/ESA SP 1284, ESA Publication Division, ESTEC, Noordwijk, The Netherlands. Available at <http://earthobservations.org>
- Gross R, Beutler G, Plag H-P (2009) Integrated scientific and societal user requirements and functional specifications for the GGOS. In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, pp 209–224
- Herring TA, Altamimi Z, Plag H-P, Poli P (2009) The future geodetic reference frame. In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, pp 225–236
- Ilk KH, Flury J, Rummel R, Schwintzer P, Bosch W, Haas C, Schroeter J, Stammer D, Zahel W, Miller H, Dietrich R, Huybrechts P, Schmeling H, Wolf D, Goetze HJ, Riegger J, Bardossy A, Guenter A, Gruber T (2005) Mass transport and mass distribution in the Earth system, Tech. rep., GOCE-Projectbuero Deutschland, Technische Universitaet Muenchen, GeoForschungsZentrum Potsdam
- Mayer-Gürr T, Ilk KH, Eicker V, Feuchtinger M, 2005. ITG-CHAMP01: a CHAMP gravity field model from short kinematic arcs over a one-year observation period. *J Geodesy* 78:462–480
- Panet I, Mikhailov V, Politz F, Diament M (2008) Insight into the Sumatra December 2004 and March 2005 post-seismic signals from GRACE gravity variations. In: Ries J, Bettadpur S (eds) *Proceedings of the GRACE Science Team Meeting, December 2008*, pp 432–438
- Plag H-P, Adegoke J, Bruno M, Christian R, Digiacomio P, McManus L, Nicholls R, van de Wal R (2010) Observations as decision support for coastal management in response to local sea level changes. Community White Paper. In: *Proceedings of OceanObs’09, September 2009, Venice, Italy*. In press
- Plag H-P, Altamimi Z, Bettadpur S, Beutler G, Beyerle G, Cazenave A, Crossley D, Donnellan A, Forsberg R, Gross R, Hinderer J, Komjathy A, Mannucci A J, Ma C, Noll C, Nothnagel A,

- Pavlis EC, Pearlman M, Poli P, Schreiber U, Senior K, Woodworth P, Zuffada C (2009a) The goals, achievements, and tools of modern geodesy. In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, pp 15–88
- Plag H-P, Beutler G, Gross R, Herring TA, Rizos C, Rothacher M, Rummel R, Sahagian D, Zumberge J (2009b) Introduction. In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, 332 pp
- Rothacher M, Beutler G, Bosch W, Donnellan A, Gross R, Hinderer J, Ma C, Pearlman M, Plag H-P, Ries J, Schuh H, Seitz F, Shum CK, Smith D, Thomas M, Velacogna E, Wahr J, Willis P, Woodworth PL (2009) The future global geodetic observing system (GGOS). In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, pp 237–272
- Rummel R, Beutler B, Dehant V, Gross R, Ilk KH, Plag H-P, Poli P, Rothacher M, Stein S, Thomas R, Woodworth PL, Zerbini S, Zlotnicki V (2009) Understanding a dynamic planet: Earth science requirements for geodesy. In: Plag H-P, Pearlman M (eds) *The global geodetic observing system: meeting the requirements of a global society on a changing planet in 2020*. Springer, Berlin, pp 89–134
- Shum CK, Braun A, Cazenave A, Chamber D, Emery W, Fotopoulos G, Gouretski V, Gross R, Miller L, Hughes C, Ishii M, Jayne S, Kuo (13) C-Y, Leuliette E, Levitus S, Maximenko N, Morison J, Plag H-P, Rothacher M, Rummel R, Schröter J, Shibuya K, Sideris M, Song YT, Willis J, Woodworth P, Zlotnicki V (2010) Geodetic observations of the ocean surface topography, geoid, currents, and changes in ocean mass and volume. Community White Paper. In: *Proceedings of OceanObs'09, September 2009, Venice, Italy*. In press
- Turner BL II, Clark WC, Kates RW, Richards JF, Mathews JT, Meyer WB (eds) (1990) *The Earth as transformed by human action: global and regional changes in the biosphere over the past 300 years*. University Press, Cambridge, 713 pp
- Watkins MM, Yuan DN, Kuang D, Bertiger W, Byun S, Lu W, Krusinga GL (2008) JPL L-2 GRACE solutions: harmonics, mascons, iteration, and constraints. In: Ries J, Bettadpur S (eds) *Proceedings of the GRACE Science Team Meeting, December 2008*, pp 82–92
- World Commission on Environment and Development (1987) *Our common future*. Oxford University Press, Oxford
- Xu P (2008) Position and velocity perturbations for the determination of geopotential from space geodetic measurements. *Celest. Mech. Dyn. Astron.* 100:231–249

Chapter 11

Monitoring Radial Tectonic Motions of Continental Borders Around the Atlantic Ocean and Regional Sea Level Changes by Space Geodetic Observations

Zhigen Yang and Fengchun Shu

Abstract In this paper, the possible radial tectonic motions of continental border of both sides of the Atlantic Ocean are estimated based on the International Terrestrial Reference Frame ITRF2000 and the NNR-NUVEL-1A plate motion model. The technique uses the observed VLBI baseline length change rates around the Atlantic Ocean as constraints to estimate the regional geometric change of the Earth, the regional GPS and SLR velocity field data are also processed respectively to confirm the above estimated results. Primarily comparison of the radial tectonic motion results with the observed mean sea level rises around the eastern coast of the North America is performed. We conclude that the average land elevation of continental border around both sides of mid-Atlantic in northern hemisphere probably has a systematically sinking of about 1.57 ± 0.35 mm/year because of the possible radial tectonic motion in this area. Therefore, the approximately same value of extra sea level rise near the eastern coast of the North America probably comes from the crustal radial tectonic motion, but the post-glacial rebound as regarded previously. Further studies and discussion to the possible radial tectonic motions mentioned above using the velocity field data in ITRF2005 and/or in ITRF2008 would be considered so as to address further the estimated results in this work.

11.1 Introduction

During the past decades, the uncertainty of the horizontal components and regional deformation solutions of Very Long Baseline Interferometry (VLBI) plate motions is about ± 1.0 mm/year. Average accuracy for global solutions of VLBI baseline rates (~ 650 global baselines or more) is better than ± 1.4 mm/year (Ma 2003; Yang et al. 2004, 2005). Studies on regional radial tectonic motions and/or deformations

Z. Yang (✉)

Shanghai Astronomical Observatory, Chinese Academy of Sciences,
80 Nandan Road, Shanghai 200030, China
e-mail: yangz@shao.ac.cn

and possible earth's geometrical changes have become possible for the intention of better understanding regional real sea level changes (Sun et al. 1999; Ma et al. 1992; Jin and Zhu 2003).

Over the past 51 year period, there is no significant difference in the rates of coastal and global mean sea level rise, as found in climate model simulations of the 20th century. The best estimate of both global and coastal sea level rise remains 1.8 ± 0.3 mm/year, as found in earlier studies (White et al. 2005). The recent estimates on the sea level rise from 1961 to 2003 is about 1.5 ± 0.4 mm/year (Domingues et al. 2008), in good agreement with their updated estimate of near-global mean sea level rise of 1.6 ± 0.2 mm/year (Church et al. 2004; Church and White 2006). However, according to the estimates from satellite altimetry since the start of the TOPEX/Poseidon (T/P) record in late 1992 (Church et al. 2004; Leuliette et al. 2004; Church et al. 2008; Beckley et al. 2007), the sea level rise near coastal areas may be near or over 3 mm/year. Satellite altimetry measurements by the TOPEX/Poseidon and Jason 1 satellite covering the years 1993–2003 provide a value of sea level rise of 3.1 ± 0.7 mm/year (IPCC 2007; Nerem et al. 2006).

Global fluctuations in sea level may result from some regional processes, such as thermal expansion of ocean waters, changes in melt water load, crustal post-glacial rebound (PGR) and uplift or subsidence in coastal areas related to various tectonic processes (Cazenave and Nerem 2004), and are also controlled by many different variables, including interactions between lithology, geomorphology and climate (Gornitz 1991), as well as more local shelf-scale changes in coastal winds, atmospheric pressures and freshwater runoff, and vertical land motions from large-scale glacial isostatic rebound, local tectonic motions and the response to changes in surface loading of the Earth from changing mass of ice sheets, glaciers and terrestrial storage (White et al. 2005).

Munk (2003) stated that global sea levels have been rising over the past century. It remains hotly debated, however, how fast this rise has been and which of the two causes: expansion of ocean waters due to warming, or freshwater input from the continents-has dominated. He cautions, however, that uncertainties in sea ice melting rates and possible biases in tide gage data do not yet allow a definitive answer. Even with global satellite coverage, several decades of data will be required for good estimates of the role of global warming in sea level rise. Miller and Douglas (2006) found that sea level trends from tide-gauges, which reflect both mass and volume change, are 2–3 times higher than rates based on hydrographic data which reveal only volume change, thereby indicating that mass increase plays a much larger role than ocean warming in twentieth century global sea level rise, which differs from the conclusion that the rate of ocean warming and thermosteric rise from 1961 to 2003 is about 50% (Domingues et al. 2008). Shepherd and Wingham (2007) further obtained that it is reasonable to conclude that, today, the East Antarctic mass is gaining some 25 Gt/year, the West Antarctic mass is losing about 50 Gt/year, and the Greenland mass is losing about 100 Gt/year, these trends provide a sea level contribution of about 0.35 mm/year, a modest component of the present rate of sea level rise of 3.0 mm/year.

However, so far we do not have a complete understanding of regional scale crustal deformation, especially the crustal radial tectonic motions. The geodetic studies on crustal vertical motion and its model solution, especially the global scale mass variations in the Earth system and a big challenge for our understanding of the geophysical fluids were presented and suggested at the third International VLBI Service for Geodesy and Astrometry (IVS) General Meeting, February 9–12, 2004, Ottawa, Canada (Fesissel-Vernier et al. 2004). VLBI/GPS/absolute gravity space geodetic measurements of crustal motions can be employed to correct many long term tide gauge records because of vertical crustal movements, improving the geographic coverage of sea level trends (Douglas 1997). Due to the accurate short/long-term records of sea level, and the mid-Atlantic coast areas in northern hemisphere can potentially be affected by vertical motion of the land as pointed above, the possible average radial tectonic motions along the continental border of mid-Atlantic region in northern hemisphere is estimated and quantified by using data of VLBI baseline length change rates (Ma 2003, 2005; Yang et al. 2004, 2005), the three dimensional (3-D) data of velocity field in rectangular coordinate system for the GPS (Heflin et al. 2004) and satellite laser ranging (SLR) (Boucher et al. 2001) measurements are analyzed to verify the above results by VLBI.

11.2 Methodology

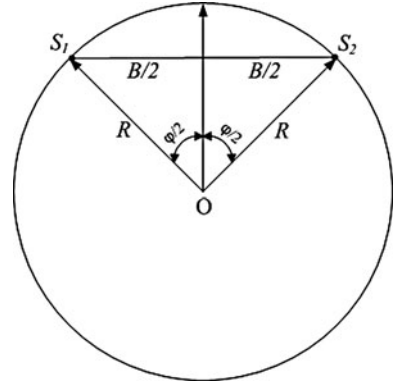
The baseline length change \dot{B} between two VLBI stations, which are located in the two different continents, can be expressed by the following components:

$$\dot{B} = R_{Mol} + R_{Reg} + R_{Loc} + R_{Geo} + R_{PGR} + R_{Rot} \quad (11.1)$$

where, the right corresponding components in Eq. (11.1), R_{Mol} is the effects from the plate motions, R_{Reg} , R_{Loc} , R_{Geo} and R_{PGR} are, respectively, from the regional crustal motions, local crustal motions, motions caused by other geophysical reasons (Gornitz 1991; Cazenave and Nerem 2004), and from the PGR, R_{Rot} represents effects from variations of earth rotation. Using the data of VLBI baseline rates, we have successfully detected and separated the difference of vertical deformation rates between collocated VLBI stations in the world (Yang et al. 2001, 2002; Yang and Zhu 2005, 2006) based on different ITRF solutions. In theory, the distance measurement and its change rate estimation along baseline between two ground VLBI stations are independent of the adopted terrestrial reference frame (Ma 2003; Yang et al. 2004, 2005). Therefore, if using the observed baseline rates as constraints to estimate possible radial tectonic motions, and considering that the effects of R_{Rot} in Eq. (11.1) can be ignored due to its smaller effects, we can well obtain a possible average radial tectonic motion rate along the continental border of mid-Atlantic region where the VLBI baselines covered in northern hemisphere.

Write B for the single-baseline length between two VLBI stations that are respectively located at continental border of both sides of the Atlantic (see Fig. 11.1), R and

Fig. 11.1 VLBI baseline B between stations S_1 and S_2 , O represents the geocentric center of the Earth



ϕ respectively, for the distance from geo-center to VLBI stations and the geocentric angle of the VLBI baseline spans, then the B can be expressed as:

$$B = 2R \sin(\phi/2) \quad (11.2)$$

Using the differential calculus to both sides of Eq. (11.2), the possible regional average radial rate span by VLBI baseline may be separated, and we could have the following equation:

$$\begin{aligned} \dot{B} &= 2\dot{R} \sin(\phi/2) + R\dot{\phi} \cos(\phi/2) \\ &= 2\dot{R} \sin(\phi/2) + \dot{\Omega} \cos(\phi/2) \end{aligned} \quad (11.3)$$

where \dot{B} is the observed baseline rate, the \dot{R} represents the average radial rate of continental zone between two VLBI stations around both sides of the Atlantic, being the unknown quantity, which need to be estimated, and $\dot{\phi}$ and $\dot{\Omega}$ are respectively the change rate of ϕ and the change rate of arc length corresponding to the VLBI baseline spans, which can be estimated from plate motion model. From Eq. (11.3), the observed baseline rate \dot{B} is composed of two parts, one is the $2\dot{R} \sin(\phi/2)$, relating with the average radial rate of VLBI stations of continental zone around both sides of the Atlantic, another one is the $\dot{\Omega} \cos(\phi/2)$, the item of horizontal motion of the VLBI stations caused by plate motion along the baseline direction of two target VLBI stations. Based on the NNR-NUVEL-1A plate motion model (DeMets et al. 1994), using the observed baseline rate \dot{B} as input constraint, it is possible to accurately estimate the average \dot{R} separated by several hundred to several thousand kilometers by Eq. (11.3), to refine the crustal regional vertical or radial tectonic motion solutions, and to improve some possible inadequacies of ITRF solutions.

11.3 Results

The adopted distribution of the nine VLBI baselines (including two baselines that are located at southern hemisphere, along the both sides of the Atlantic in northern hemisphere is shown in Fig. 11.2, the baseline length change rates and its corresponding results of radial rate estimate are respectively shown in Table 11.1.

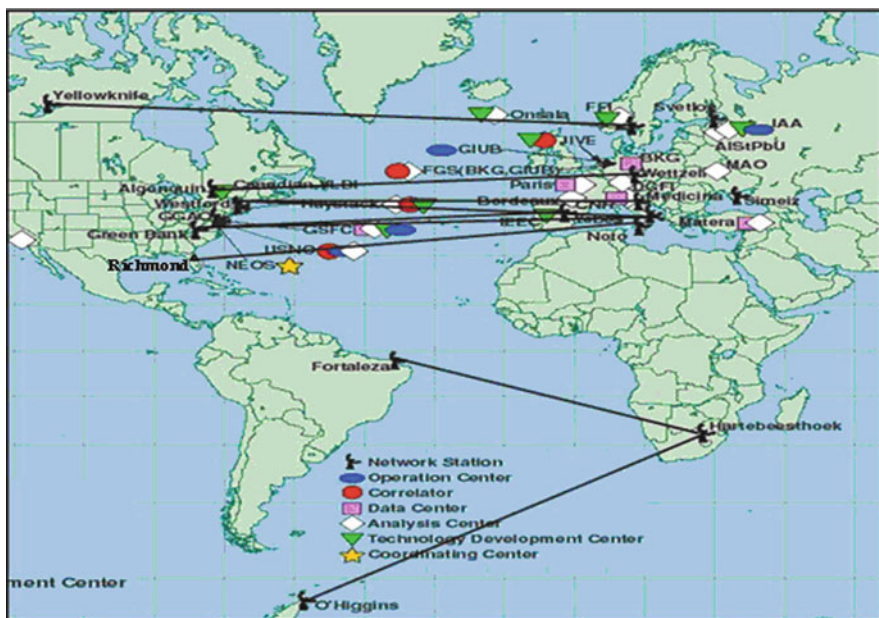


Fig. 11.2 Distribution of nine VLBI baselines along the both sides of east-westward continental border of the Atlantic. The original figure is from IVS publication (Boucher et al. 2001)

Table 11.1 The \dot{R} results from VLBI measurements along different latitude zone of mid-Atlantic in northern hemisphere: PGR of ICE-5G (Peltier 2004) is removed and PGR is not removed in bracket

VLBI baseline	Baseline rate (mm/year)	Latitude (°)	\dot{R} (mm/year)
O'Higgins-HartRAO	11.5 ± 1.3	-45.1	$-2.35(-2.97) \pm 1.21$
Fortleze-HartRAO	18.8 ± 0.4	-14.9	$-4.11(-4.53) \pm 0.36$
Richmond-NOTO	12.4 ± 3.8	+31.2	$-3.56(-3.35) \pm 3.04$
NRAO20-Yebes	21.9 ± 4.1	+39.5	$+2.03(2.37) \pm 4.28$
NRAO85_3-Matera	13.7 ± 1.6	+39.6	$-4.50(-4.03) \pm 1.39$
Westford-Matera	16.0 ± 0.2	+41.7	$-3.39(-3.08) \pm 0.29$
Westford-Medicina	15.7 ± 0.4	+43.6	$-3.85(-3.75) \pm 0.42$
Algotpark-Wettzell	18.8 ± 0.1	+47.7	$-1.21(-0.86) \pm 0.10$
Ylow7296-Onsala60	15.7 ± 2.4	+59.9	$-1.67(-2.20) \pm 2.68$

In order to verify the above estimated results, we have used the same method to calculate the radial tectonic motion rates by using the average 3-D regional velocity field data of GPS and SLR in rectangular coordinate system (Boucher et al. 2001; Heflin et al. 2004), the corresponding estimated results are shown in Table 11.2 and 11.3, respectively. In which, the \dot{R} results with removing and no removing effects of PGR are respectively listed.

In the estimation above, the GPS and/or SLR baseline length change rates were respectively estimated with weighted average of 3-D velocity field data of multi-stations (see first column of Tables 11.2 and 11.3), the effects from regional and local crustal motions, motions caused by other geophysical reasons, and part of PGR etc. to the original velocity field solution of single GPS or SLR station have been in fact treated as random errors to a large extent in the regional average velocity field estimation of GPS and/or SLR, and also in the regional average baseline rate solutions of GPS and/or SLR, which probably have produced relatively larger uncertainties for the results of \dot{R} in Tables 11.2 and 11.3. However, the estimated radial rate results in Tables 11.1–11.3 seem to be good consistent with each other.

The average and weighted average results of \dot{R} for different space techniques in Tables 11.1–11.3 are respectively given in Table 11.4. In the weighted estimation above, the weight W_p is expressed as:

Table 11.2 The \dot{R} results from GPS measurements along different latitude zone of mid-Atlantic in northern hemisphere: PGR of ICE-5G removed and PGR is not removed in bracket

GPS baseline (GPS Station number)	Baseline rate (mm/year)	Latitude (°)	\dot{R} (mm/year)
SOAM(5)-AFRC(3)	26.26 ± 3.23	-31.6	$-2.98(-3.42) \pm 3.40$
SOAM(3)-AFRC(5)	17.69 ± 11.36	-7.7	$-6.38(-6.76) \pm 9.30$
CARB(3)-AFRC(5)	-1.82 ± 11.85	+4.7	$-5.30(-5.62) \pm 7.81$
NOAM(4)-EURA(4)	14.79 ± 6.98	+38.1	$-3.43(-2.85) \pm 5.69$
NOAM(4)-EURA(25)	19.61 ± 2.23	+45.2	$-0.35(-0.53) \pm 2.53$
NOAM(2)-EURA(9)	16.19 ± 3.09	+69.1	$+0.38(-1.70) \pm 6.85$

Table 11.3 The \dot{R} results from SLR measurements along different latitude zone of mid-Atlantic in northern hemisphere: PGR of ICE-5G removed and PGR is not removed in bracket

SLR baseline (SLR Station number)	Baseline rate (mm/year)	Latitude (°)	\dot{R} (mm/year)
SOAM(3)-AFRC(2)	15.65 ± 3.97	-28.4	$-6.92(-7.57) \pm 3.06$
SOAM(3)-AFRC(3)	12.89 ± 3.73	+33.2	$-3.20(-2.82) \pm 3.07$
NOAM(3)-EURA(7)	19.71 ± 7.48	+34.8	$+0.39(+0.87) \pm 5.36$
NOAM(2)-EURA(6)	18.45 ± 2.91	+44.8	$-1.79(-1.43) \pm 2.99$
NOAM(2)-EURA(2)	16.19 ± 3.09	+54.6	$+0.26(-1.98) \pm 4.92$

Table 11.4 Radial tectonic motions around middle latitude zone of the North Atlantic by VLBI, GPS and SLR, respectively. PGR of ICE-5G removed and PGR is not removed in bracket

Techniques	Average \dot{R} (mm/year)	Weighted average \dot{R} (mm/year)
VLBI	$-2.31 \pm 0.86(-2.12 \pm 0.85)$	$-1.57 \pm 0.35(-1.24 \pm 0.37)$
GPS	$-2.17 \pm 1.33(-2.67 \pm 1.09)$	$-1.06 \pm 0.93(-1.32 \pm 0.84)$
SLR	$-1.08 \pm 0.86(-1.34 \pm 0.79)$	$-1.75 \pm 0.56(-1.73 \pm 0.65)$

$$W_p = (\sigma_i)^{-2} \tag{11.4}$$

Here σ_i is the error of \dot{R} results in Tables 11.1, 11.2 and 11.3, respectively. Thus, the adopted data number of \dot{R} results for calculating the weighted average of the different techniques is 7 for VLBI, 4 for GPS and SLR, respectively. In Table 11.4, if we take the weighted average of \dot{R} results of GPS and SLR with PGR were removed in the third column, we then have $\dot{R}_{(GPS + SLR)} = -1.57 \pm 0.30$ mm/year, which is almost the same value with that of the VLBI in the third column of the Table 11.4. This means that around the continental border area of both sides of the Atlantic, the radial tectonic motions is much probably in existence, which would has been included in the observed data of sea level rise.

11.4 Sea Level Rise

Holgate and Woodworth (2004) found evidence that the coastal sea level was rising significantly faster than global mean sea level. The Goddard Institute for Space Studies (2004) had the result that sea level in New York City has risen from 2.3 to 3.8 mm/year over the last hundred years, due to the major factors including thermal expansion as the ocean warming, melting of mountain glaciers and subsidence (sinking of the East Coast due to isostatic adjustments of the crust from the last ice age). At the coastal waters of Washington State, local sea level rise is produced by the combined effects of global sea level rise and local factors such as vertical land deformation (Mote et al. 2008).

Additionally, based on the report of Titus (2002), the sea level has already been rising along the U.S. coast. IPCC (2001) estimated that global sea level will rise 9–88 cm during the 21st century. Furthermore, due to the regional subsidence, the rise will be 15–25 cm greater in parts of the mid-Atlantic, and 5–15 cm greater elsewhere along the Atlantic Coast, i.e. averagely, during the 21st century the sea level rise of about 6.8 mm/year and 5.8 mm/year will occur respectively in parts of mid-Atlantic and at elsewhere along the Atlantic Coast. According to Titus (2002), considering the effects both subsidence and greenhouse gases, sea level is most likely to rise by about 2 feet along most of the Atlantic Coast, that means the sea level rise will be about 6.2 mm/year in 21st century around most of the

Atlantic Coast area, which would threaten most wetlands along the Eastern Shore of Chesapeake Bay, and a large number of coastal wetlands in the mid-Atlantic Titus et al. (2009). It should be noted that the regional subsidence around both sides of continental zone of the north Atlantic would be about 1.5–2.5 mm/year as mentioned by Titus (2002).

The interval 1930–1980 is long enough at 50 years to establish that the mid-Atlantic region has a systematically higher rate of sea level rise than the long term global average of nearly 1–2 mm/year (Fig. 11.3). The approximately 1.5 mm/year extra rise for the region comes from a possible sinking of the land, due to the regional radial tectonic motions (VanDam and Schenewerk 1997). A similar plot just for tide gauges along the east coast of North America and tide gauges on the Chesapeake Bay is emphasized using a different color. Note the similarity of all these plots.

Figure 11.4 shows, clearly, the linear trend result of 2.9 mm/year, for the entire 103-year span (black line) from 1893 to 1995, of the annual mean relative sea level at New York City (Douglas and Peltier 2002). Rosenzweig and Solecki (2001) concluded that the sea level in New York City has risen on average 2.7 mm/year or 2.3–3.8 mm/year over the last hundred years. Looking ahead, it is expected that sea level in the area will rise on average 3.9 mm/year or anywhere from 1.8 to 6.0 mm/year by the 2050. The researchers projected a rise in sea level of 11.8–37.5 inches in New York City and 9.5–42.5 inches in the metropolitan region by the 2080s (NASA/GSFC 2006).

Barbosa and Silva (2009) analyzed the long-term sea level variability in Chesapeake Bay, concluded that the linear trends in relative sea level rise range from 2.66 ± 0.075 mm/year (at Baltimore) to 4.40 ± 0.086 mm/year (at Hampton Roads) for the 1955–2007 period, confirming again a extra rise of sea level in the region of Chesapeake Bay.

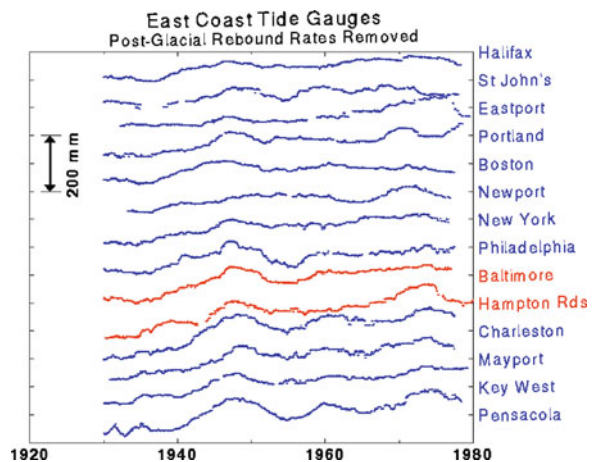


Fig. 11.3 Sea level changes of stations along U.S. eastern seaboard by the east coast tide gauges from 1930 to 1980 (ICE-3G PGR removed) (Douglas and Lilibridge 1994)

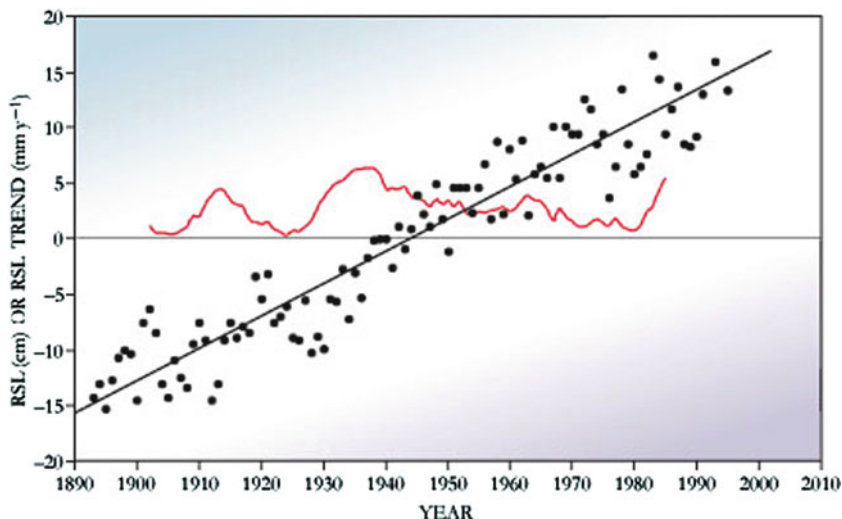


Fig. 11.4 Mean sea level changes of New York City from 1893 to 1995. *Black dots* represent the annual relative sea level measured in centimeters. The *red line* represents the running 20-year mean rate of relative sea level measured in millimeters, in which the interannual variations are clear (see Douglas and Peltier 2002)

11.5 Conclusion and Discussion

The North Atlantic, including its east-westward continental border would have a systematically crustal radial subsidence as 1.57 ± 0.35 mm/year based on the multi-space geodetic observations (Table 11.4), which is consistent with the estimates by many authors, such as Titus (2002). The baseline rate measurements made by space geodetic observing techniques, especially the VLBI, can potentially be used as constraints to detect and/or separate some signals, such as the regional and/or large scale area crustal radial tectonic motions.

In the mean sea level rise of 3.0 mm/year or over mentioned above, about 1.6 ± 0.3 mm/year excessive rate of sea level rise at east coastal area of North America can probably be interpreted by regional radial tectonic subsidence. Therefore, the true sea level rise in this area should be about ~ 1.5 mm/year, which includes the possible less effects of thermal expansion to the sea level rise as the ocean warming (Miller and Douglas 2006; Munk 2003). According to the estimate by Shepherd and Wingham (2007), the Antarctica and Greenland are each losing mass overall, which, as the best estimate, is causing sea level rise by 0.35 mm/year, is only a modest contribution to the present rate of sea level rise of 3.0 mm/year. In consequence, the view that the changing sea level contribution of the Antarctic and Greenland ice sheets in the 21st century will be both small and negative as a result of accumulating snow in Antarctica (e.g., -0.05 mm year⁻¹) is now uncertain, because our predictive ability is limited, continued observation is essential. Church et al.

(2004) confirmed that a greater rate of sea level rise on the eastern North American coast compared with the United Kingdom and the Scandinavian Peninsula is also found.

Over the 51 year period from 1950 to 2000, there is no significant difference in the rates of coastal and global averaged sea level rise, as found in climate model simulations of the 20th century. The best estimate of both global average and coastal sea level rise remains 1.8 ± 0.3 mm/year, as found in earlier studies. If extend the reconstruction of global mean sea level back to 1870, it is found a sea level rise from January 1870 to December 2004 of 195 mm, a 20th century rate of sea level rise of 1.7 ± 0.3 mm/year and a significant acceleration of sea level rise of 0.013 ± 0.006 mm year⁻² (Church and White 2006). Domingues et al. (2008) found that the sum of contributions to sea level rise from 1961 to 2003 is about 1.5 ± 0.4 mm/year, in good agreement with their updated estimate of near-global mean sea level rise of 1.6 ± 0.2 mm/year (Church and White 2006). However, Cazenave and Nerem (2004), and Leuliette et al. (2004) concluded that the geocentric rate of global mean sea level rise over 1993–2003 is 2.8 ± 0.4 mm/year by T/P and Jason altimeter data, 3.1 mm/year if PGR is removed.

Our result on crustal radial subsidence along the east-westward continental border of the North Atlantic seems consistent with the conclusion that the north hemisphere is a compressive hemisphere derived by space geodesy (Huang et al. 2002; Jin and Zhu 2003; Sun et al. 1999) because of the observed results of negative closed difference of plate loop measured by space geodetic observations, which suggests that the latitude circle is shortening. The previous discussions concluded that the north hemisphere is a cold and compressing hemisphere, while the south hemisphere is a hot and expanding hemisphere based on analyzing the data of dissipation of earth's heat flows, the distribution of seismic waves, and data of the magnetic strips of global ocean floors (Fu 1976; Ma et al. 1992). Additional, our result potentially validated (see Tables 11.1, 11.2, and 11.3) the conclusion that at the north hemisphere the contractive rate gradually decreases from the equator to the North Polar area (Jin and Zhu 2003).

Therefore, we conclude that the systematically crustal radial subsidence may be in existence in some regions, such as around north Atlantic area. The similar results of modern-day tectonic subsidence in coastal Louisiana were obtained by Dokka (2006). According to Miller and Douglas (2007) and Gehrels et al. (2006), and the long tide gauge records in the North Atlantic and North Pacific, it was shown that little increase in sea level before 1930, and followed by mostly a steady rise from 1930. In addition, the land subsidence associated with materials extraction and/or underground exploitation, such as petroleum and natural gas extraction, probably was one of potential factors. Therefore, scientists believe that the sea level rises could far exceeded IPCC estimates (Brahic 2008), the discussion on sea level rise in 20th century and recent few decades is far from being closed.

The International Terrestrial Reference Frame ITRF2005 (Altamimi et al. 2007) was issued in 2006, and the ITRF2008 will be issued in 2009–2010, in which the VLBI solution will be spanned from 1980 to 2009, and GPS and SLR preliminary solutions will respectively extended backward from 1997 to 2009 and from 1984 to

2009. Therefore, further choosing and adopting the data of collocated sites including the VLBI baseline rate measurements, 3-D data of GPS and SLR observations as constraints to re-estimate and discuss the possible radial tectonic motions mentioned above would be necessary, so as to really confirm and better understanding the regional radial tectonic motions along the continental border of the Atlantic, although in this studies the regional average of data of the GPS and SLR 3-D velocity fields were used, in which the regional and local crustal motions, motions caused by other geophysical reasons, and part of PGR can be treated to some extent as random errors, which probably produced relatively larger uncertainties for the results of \dot{R} as seen in Tables 11.2 and 11.3. Besides, the DORIS data reductions can also be used for the future discussion of possible regional radial tectonic motion solution.

Acknowledgments This work is supported by key laboratory of dynamic geodesy and geophysics, Chinese academy of sciences (L06-03), Shanghai Natural Science Foundation (09ZR1437300) and Key Technologies and Application Research and Develop of Shanghai (08DZ1500300). Prof. Ye Shuhua reviewed the manuscript. The authors would like to thank Dr. Junathan Li of the University of Waterloo, Canada and other reviewers for their reviews and helpful suggestions on this manuscript. The paperwork of this research was mainly completed at Shanghai Galileo Industries, Ltd.

References

- Altamimi Z, Collilieux X, Legrand J, Garayt B, Boucher C (2007) ITRF2005: a new release of the International Terrestrial Reference Frame based on time series of station positions and Earth orientation parameters. *J Geophys Res* 112:B09401. doi:10.1029/2007JB004949
- Barbosa SM, Silva ME (2009) Low-frequency sea level change in Chesapeake Bay: changing seasonality and long-term trends, *Estuarine, Coastal Shelf Sci* 83(1):30–38
- Beckley BD, Lemoine FG, Lutcke SB, Ray RD, Zelensky NP (2007) A reassessment of global and regional mean sea level trends from TOPEX and Jason-1 altimetry based on revised reference frame and orbits. *Geophys Res Lett* 34:L14608
- Boucher C, Altamimi Z, Sillard P et al. (2001) IERS Technical Note No. 31, Observatories de Paris Brahic C (September 2008) Sea level rises could far exceeded IPCC estimates, *New Scientist, Environment*. 12:41
- Cazenave A, Nerem RS (2004) Present-day sea level change: observations and causes. *Rev Geophys* 42:RG3001
- Church JA, White NJ, Coleman R, Lambeck K, Mitrovica JX (2004) Estimates of the regional distribution of sea level rise over the 1950 to 2000 period. *J Clim* 17:2609–2625
- Church JA, White NJ (2006) A 20th century acceleration in global sea level rise. *Geophys Res Lett* 33:L01602
- Church JA, White NJ, Aarup T, Wilson WS, Woodworth PL, Domingues CM, Hunter JR, Lambeck K (2008) Understanding global sea levels: past, present and future. *Sustainability Science* 3(1):9–22
- DeMets C, Gordon R, Argus D et al. (1994) Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. *Geophys Res Lett* 21:2191–2194
- Dokka RK (2006) Modern-day tectonic subsidence in coastal Louisiana. *Geology* 34(4):281–284
- Domingues CM, Church JA, White NJ, Gleckler PJ, Wijffels SE, Barker PM, Dunn JR (2008) Improved estimates of upper-ocean warming and multi-decadal sea level rise. *Nature* 453:1090–1094
- Douglas BC (1997) Global sea rise: a redetermination. *Surv Geophys* 18(2–3):279–292

- Douglas B, Lilibridge J (1994) Variations in sea level. <http://www.ngs.noaa.gov/GRD/GPS/Projects/CB/SEALEVEL/sealevel.html>. Accessed on June 2010
- Douglas BC, Peltier WR (2002) The puzzle of global sea level rise. *Phys Today* 35–40
- Fesissel-Vernier M, Ray J, Altamimi Z, Dehant V, de Viron O (2004) VLBI and earth rotation: geophysical and geodetic challenges, General Meeting Proceedings 2004, International VLBI Service for Geodetic and Astrometry, Ottawa, Canada, February 9–11, pp 22–31
- Fu C (1976) Tenth lecture of the Earth. Science Press, Beijing, pp102–159
- Gehrels WR, Marshall WA, Gehrels MJ, Larsen G, Kirby JR, Eiríksson J, Heinemeier J, Shimmield T (2006) Rapid sea level rise in the North Atlantic Ocean since the first half of the nineteenth century. *The Holocene* 16(7):949–965
- Goddard Institute for Space Studies (2004) Research results on climate impacts in New York City: sea level rise and coastal floods, NASA Official. <http://icp.giss.nasa.gov/research/ppa/2002/impacts/introduction.html>. Accessed on June 2010
- Gornitz V (1991) Global coastal hazards from future sea level rise, palaeogeography, palaeoclimatology, palaeoecology. *Global Planetary Change Sect* 89:379–398
- Heflin et al. (2004) GPS Time Series, JPL, California Institute of Technology. <http://sideshow.jpl.nasa.gov/mbh/series.html>. Accessed on June 2010
- Holgate SJ, Woodworth PL (2004) Evidence for enhanced coastal sea level rise during the 1990's. *Geophys Res Lett* 31:L07305
- Huang L, Ma Z, Zhu J (2002) New measurement evidence of earth asymmetry. *Acta Seismol Sinica* 15(2):210–213
- IPCC (2001) The IPCC Third Assessment Report “Climate Change 2001”. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds.). Cambridge, UK and New York, NY, USA
- IPCC (2007) In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.) *Climate change 2007: the physical science basis, contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- Jin S, Zhu W (2003) Compression of the north hemisphere derived from space geodesy. *Acta Seismol Sinica* 16(1):99–106
- Leuliette EW, Nerem RS, Mitchum GT (2004) Calibration of TOPEX/Poseidon and Jason altimeter data to construct a continuous record of mean sea level change. *Marine Geodesy* 27(1&2): 79–94
- Ma C (2003, 2005) On glb2003cn baseline rate solutions. VLBI Group at NASA Goddard Space Flight Center (private communications)
- Ma Z, Gao X, Ren J (1992) Present day explain of tectonic character and dynamics. *Quaternary Res* (4):293–305
- Miller L, Douglas BC (2006) On the rate and causes of twentieth century sea level rise. *Phil Trans R Soc A* 364(1841):805–820
- Miller L, Douglas BC (2007) Gyre-scale atmospheric pressure variations and their relation to 19th and 20th century sea level rise. *Geophys Res Lett* 34(L16602):1–5
- Mote P, Petersen A, Reeder S, Shipman H, Binder LW (2008) A report by the University of Washington Climate Impacts Group and the Washington Department of Ecology on sea level rise in the coastal waters of Washington State. <http://cses.washington.edu/db/pdf/moteetalslr579.pdf>. Accessed on June 2010
- Munk W (2003) Ocean freshening, sea level rising. *Science* 300(5628):2041–2043
- NASA/Goddard Space Flight Center (October 26, 2006) NASA looks sea level rise. Hurricane risks to New York City
- Nerem R, Leuliette E, Cazenave A (2006) Present day sea level change: a review. *CR Geosci* 338:1077–1083
- Peltier WR (2004) Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G (VM2) model and GRACE. *Ann Rev Earth Planet Sci* 32:111–149

- Rosenzweig C, Solecki WD (eds) (2001) Climate change and a global city. Columbia Earth Institute, New York
- Shepherd A, Wingham D (2007) Recent sea level contributions of the Antarctic and Greenland ice sheets. *Science* 315(5818):1529–1532
- Sun F, Zhao M, Ning J et al (1999) Detection of unsymmetrical global tectonic change by using space geodetic data. *Chin Sci Bull* 44(20):2225–2229
- Titus JG et al. (2009) U.S. Climate Change Science Program and the Subcommittee on Global Change Research. Final Report Synthesis and Assessment Product 4.1, Coastal sensitivity to sea level rise: a focus on the Mid-Atlantic region
- Titus J (2002) Does sea level rise matter to transportation along the Atlantic coast? The potential impacts of climate change on transportation, Workshop, October 1–2, 2002, Summary and Discussion Papers, U.S. Department of Transportation, pp 1–16
- VanDam T, Schenewerk M (1997) Variations in Sea Level, Chesapeake Bay project <http://www.ngs.noaa.gov/GRD/grdprojects.shtml>. Accessed on June 2010
- White NJ, Church JA, Gregory JM (2005) Coastal and global averaged sea level rise for 1950 to 2000. *Geophys Res Lett* 32:L01601
- Yang Z, Berube M, Searle A (2004) Relative deformations between co-located VLBI stations and comparisons with VTRF2003. In: Richter B, Dick WR, Schwegmann W (eds) *Proceeding of the IVS 2004 General Meeting*, Ottawa, Canada, NASA/CP-2004-212255, February 9–11, pp 481–485
- Yang Z, Berube M, Searle A (2005) Relative deformation rates of collocated VLBI stations and comparisons with global solutions. *Chin Astronomy Astrophys* 29:303–308
- Yang Z, Shu F (2006) Relative deformation rates between collocated VLBI stations and its comparisons with VTRF2005. *Annals of Shanghai Astronomical Observatory, Chinese Academy of Sciences* 27:26–32
- Yang Z, Zhu Y (2005) Estimations of vertical deformation rates of VLBI stations by constraints of rates of baseline length change. *Ann Shanghai Observatory, Acad Sinica* 26:7–13
- Yang Z, Zhu W, Cheng Z (2001) Determination and discussion of motion of Shanghai VLBI station relative to Eurasian plate. *Acta Geod Cartograph Sinica* 30(1):10–15
- Yang Z, Zhu W, Shum CK, Shu F (2002) Determination of the regional deformation rates of Shanghai and Kashima VLBI stations based on ITRF97. *Chin Astronomy Astrophys* 26:497–502

Chapter 12

GNSS Activities for Natural Disaster Monitoring and Climate Change Detection at GFZ – An Overview

Junping Chen, Michael Bender, Geory Beyerle, Galina Dick, Carsten Falck, Maorong Ge, Gerd Gendt, Stefen Heise, Marbus Ramatschi, Torsten Schmidt, Ralf Stosius, and Jens Wickert

Abstract Natural hazards and climate change are of major concern to the society. Huge losses are reported in recent years. It is widely believed that modern GNSS technologies are effective in hazard monitoring and climate change detection and modelling. Considering the limitations of current accuracy and reliability, sophisticated strategies and models have to be developed. We introduce and overview recent GNSS activities at GFZ in this field, including ground and satellite based atmospheric sounding, reflectometry and GNSS seismology. In addition we summarize recent hardware developments, where new robust on-site hardware systems combining GNSS receiver and other sensors (e.g. seismic sensors and weather sensors) are developed. The main focus of our contribution is recent results of GNSS analysis software developments for real-time applications, where multi-technique (e.g. SLR and GNSS) and multi-system (e.g. GPS and GLONASS) data source can be handled uniquely. The software can run in real-time as well as post-processing modes, and precision of several mm for ground surface deformation can be achieved. We overview the ground and spaced based GNSS atmosphere researches based on the estimation and assimilation of atmosphere parameters, which are among those parameters estimated from our GNSS software. Related projects, applying these new developments, are also introduced.

12.1 Introduction

Natural hazards are of major concern to the society. DMISCO (Disaster Management International Space Coordination Organization) stated: From 1994 to 2003 there were more than 300 major natural disasters on average each year, impacting more than 100 countries, killing over 50,000 people, affecting nearly

J. Chen (✉)
Helmholtz Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, 14473 Potsdam, Germany
e-mail: junping.chen@gfz-potsdam.de

260 million people and causing economic damage totaling US\$ 55 billion each year (GeoForschungsZentrum Potsdam 2006). In 2008, the Wenchuan earthquake in China caused 69,225 known deaths, 17,939 people are listed as missing, and 374,640 injured (Xinhua News Agency 2009).

For more than 10 years now, global and regional GPS/GNSS (Global Navigation Satellite System) networks have demonstrated their potential for detecting plate tectonic movements, strain and stress accumulation. During the mega-earthquake off Sumatra in December 2004 a dense network of GPS stations in Thailand demonstrated the capability for the detection and the closed monitoring of dynamic processes during a rupture process (GeoForschungsZentrum Potsdam 2006).

To achieve the real-time deformation detecting and monitoring with GNSS, the following tasks are most challenging:

- the design and development of a new series of multi-parameter stations running unattended and autonomously for long periods,
- the development of new and automated GPS/GNSS software systems allowing real-time data analysis and self-detection of events, and
- the development and test of new data communication strategies for high-rate and high-volume data.

We present the latest progress at GFZ in the above tasks. Sections 12.2 and 12.3 introduce the developments of sensor stations and its application in Indian Ocean Tsunami early warning system. Section 12.4 shows the progress of GNSS reflectometry for surface deformation monitoring. The structure and applications of the newly designed GNSS software are discussed in Section 12.5.

Climate changes can also be monitored and detected by GNSS technologies. Crossing the atmosphere, GNSS signals experience propagation, which can be modelled by atmosphere parameters. These parameters are among the various estimated parameters of GNSS data analysis and are assimilated in atmosphere modelling thereafter. Recent years we see big progresses of atmosphere modelling in GNSS meteorology and atmospheric sounding, which improve our understanding of climate mechanism enormously. We summarize our climate change monitoring related activities in ground and space based GNSS atmospheric sounding in Section 12.6.

12.2 GNSS Sensor Station Developments

The geodetic branch (Department 1) at GFZ has a strong background in the field of geodetic sensor system operation such as SLR (Satellite Laser Ranging), PRARE (Precise Range And Range Rate Equipment) and instruments on satellites (e.g. onboard ERS-2, CHAMP, GRACE, TerraSAR-X). This includes the operation of sensor networks and the development of task and environment adapted sensor stations with small series manufacturing. Some examples for GFZ-developed sensor

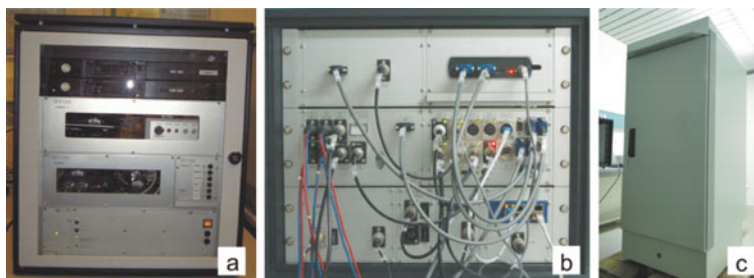


Fig. 12.1 (a) GNSS sensor station used for GFZ-operated global network locations, following a “no moving parts” and “off the shelf” components design philosophy. (b) GNSS sensor station with low power consumption and service friendly design developed for the GITEWS project. (c) Outdoor box with special passive cooling design

stations are shown on Fig. 12.1. All mentioned station designs have been tested in a climate chamber to ensure a reliable performance under all considerable climate conditions.

The station type on Fig. 12.1a was developed for the globally distributed, GFZ-operated GNSS sensor station network. It reflects findings from many years of global network operation with respect to remote hardware control, monitoring, failure recovery and repairs at remote locations. Nearly all integrated devices are “off the shelf” components with standard dimensions (Falck et al. 2008). They have been selected for long term operation but may be easily replaced after years by “off the shelf” components of a later generation. The construction has no moving parts (e.g., solid state hard disks, no fans) and produces no noise. An UPS (uninterruptible power supply) is integrated as well as a fold-out keyboard, mouse and monitor panel. Advanced remote monitoring and control devices allow failure detection and recovery, largely independent from local operators. The resulting key feature of this station type is the high level of reliability and redundancy.

Another example of a GFZ developed GNSS station design is the RTR (real-time reference) station as used for the GNSS component of the Indonesian Tsunami Early Warning System INATEWS (see next chapter). Special requirements from the field of operation are taken into account. Most important was to support the easiness of low level service works and the capability to handle frequent mains power outages enduring from seconds to days. The electronic components are installed in service-friendly modules, connected through front side cables. Even “off the shelf” components have their own extra cases to allow easy replacement in case of a malfunction. A set of rechargeable batteries is placed next to the indoor rack or inside the outdoor rack and allows a system operation of more than 2 days, independent of mains power.

Both station types (Fig. 12.1a and b) are based on a 19-inch rack mount construction housed in either regular indoor or weatherproof outdoor racks (Fig. 12.1c). The selected outdoor rack is a double wall construction, allowing passive cooling (no fans) without any kind of open window between electronics and environment. The

outdoor rack was not developed at GFZ, but identified as an important component for long term field installations.

A special software package with a modular concept has been developed. It was initially used for global network stations and adapted for the GITEWS RTR stations. The software supports several GNSS receiver types of various manufacturers as well as different meteorological sensor stations. It is capable of combining real time data streaming with file based transfer to keep the communication bandwidth low. An integrated monitor checks the status of the sensor station regularly. Automatic software updates are possible without interruption of normal operation. This software package is an essential precondition for near real-time and real-time applications.

12.3 GNSS-Based Component for Tsunami Early Warning Systems

GNSS technologies offer a high potential to support tsunami early warning systems. After the Tsunami event of 26th December 2004 the German government initiated the GITEWS project (German Indonesian Tsunami Early Warning System) to develop a tsunami early warning system for Indonesia. The new developed GNSS-based component utilises on- and off-shore measured GNSS data and is the first system of its kind that was integrated into an operational early warning system. Figure 12.2 shows the GNSS-component operator desk at the Indonesian Tsunami Early Warning Centre INATEWS at BMKG Jakarta, inaugurated on November, 11th 2008.

The new GNSS-based system covers all aspects from development, manufacturing and installation of sensor station hardware, real-time data transfer issues, a new developed automatic near real-time data processing and a graphical user interface for early warning centre operators including training on the system (Ramatschi et al. 2008). GNSS sensors are installed on buoys, at tide gauges and as real-time RTR stations, either stand-alone or co-located with seismological sensors. The GNSS data are transmitted via satellite links to the warning centre, where they are processed in a near real-time data processing chain.



Fig. 12.2 GNSS operator desk at the Indonesian Tsunami Early Warning Centre

There are two modes of data processing, the normal mode and the tsunami mode. The latter is selected as soon as a potentially tsunami relevant event was detected by the warning centre (strong earthquake, increased sea level at tide gauge, etc.). The first step in tsunami mode is the processing of data from the real-time reference stations. Then the system processes data of the 10 most relevant sensors (e.g. located nearest to the earthquake location) in 2 min intervals. In normal mode data of all sensors are processed in 5 min intervals, to allow a continuous sensor performance monitoring.

For sensors on land the data processing delivers deviations from the sensor location mean coordinates. Deviations with significant higher values than the normal noise level are regarded as land mass movements which can occur, e.g., due to strong earthquakes. This ground motion (plate tectonics) information is a valuable source for a fast understanding of an earthquake's mechanism with possible relevance for a potential following tsunami. It is also important to know displacements for locations with tide gauge sensors to separate real, tsunami caused sea level changes from apparent, displacement caused sea level changes which moved the tide gauge sensor.

For GNSS data measured on a buoy the processing (single baseline solution with one on land GNSS station as reference) delivers coordinates as well. Only the vertical component is of interest as it corresponds to instant sea level heights. Deviations to the mean sea level height are an indicator for a passing tsunami wave. By this means, ground motion and sea level height monitoring, the GNSS system supports the decision finding process whether most probably a tsunami has been generated or not.

The GUI (graphical user interface, see displays on Fig. 12.2) of the GNSS-based system supports both, a quick view for all staff members at the warning centre (24 h/7d shifts) and deeper analysis by GNSS experts. The GNSS GUI system is implemented as a web-based application and allows all views to be displayed on different screens at the same time, even outside the warning centre. This is part of the concept and supports the teamwork between warning centre staff on duty or on standby and sensor station maintenance staff.

12.4 GNSS Reflectometry

GNSS reflectometry (GNSS-R) is a promising new approach proposed by Martín-Neira (1993) that uses GNSS signals reflected from the earth to derive information about the height and the condition of the reflecting surface. It can be used in an altimetric and a scatterometric manner to measure surface height as well as wind speed, wind direction, soil moisture or sea ice extend, among others. The usability of GNSS-R has been demonstrated in several ground-based, air-borne and space-borne experiments. Code based and phase based GNSS-R approaches have to be distinguished. In code based GNSS-R the reception delay between the direct and the reflected GNSS signal is measured, which is similar to radar altimetry. Additionally the reflected signal contains information about the scattering characteristic, which

can be related to surface roughness. For the phase based approach coherent reflections are necessary. These can be obtained at grazing angle geometries, which are common also in radio occultation measurements (Beyerle and Hocke 2001). The phase interference between direct and reflected GNSS signals can be interpreted as height variation (Helm 2008). Compared to conventional measurement techniques GNSS-R has a variety of advantages. In contrast to monostatic methods like radar altimetry a GNSS-R receiver aboard a low earth orbiting (LEO) satellite receives GNSS reflections from many directions simultaneously so that GNSS-R can be regarded as multistatic (Fig. 12.3). This results in an increase of spatial and temporal resolution needed to observe mesoscale features. The GNSS signals are freely available and are used as signals of opportunity. Considering the planned installation of Galileo and forthcoming systems the number of GNSS satellites and therefore the number of reflected signals is going to increase dramatically within the next decade. Their availability will be continuous over a long time, because they are used commercially. GNSS-R is a passive measurement technique with a low energy budget. This allows the building of small and affordable GNSS-R satellites, especially when commercial-off-the-shelf (COTS) GNSS receivers could be used. The GFZ has carried out ground-based experiments with a modified COTS JAVAD receiver called GNSS Occultation, Reflectometry and Scatterometry (GORS) receiver and has demonstrated its capability to measure lake surface height at centimeter accuracy (Helm et al. 2008). When installed on small satellites within a constellation this technique is believed to be applicable as a global tsunami early warning system. A feasibility study of GFZ shows that a Sumatra like tsunami (Fig. 12.3) would be

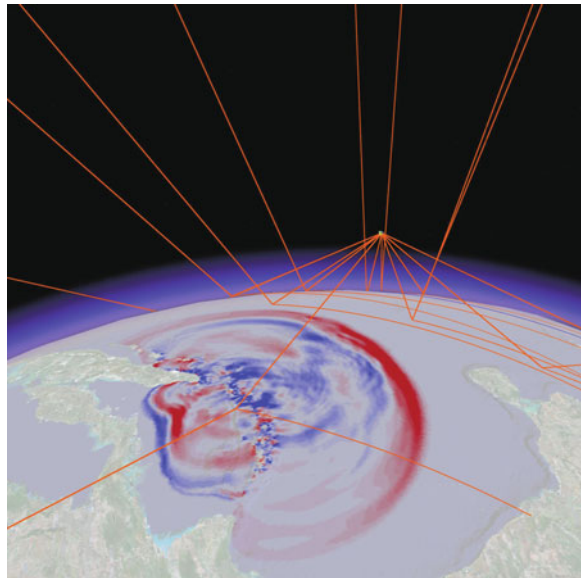


Fig. 12.3 Simulation of GNSS-R signal paths and reflection tracks during a tsunami event off-shore Indonesia

detectable within 17 min with a GNSS-R Walker-constellation of 18 satellites at 900 km altitude and 60° inclination (Stosius et al. 2008).

12.5 GNSS Seismology

Based on the expertise on GNSS software development, a newly-designed software package, EPOS-RT (Earth satellite Positioning and Orbit determination System in Real-Time) (Chen et al. 2008, 2009; Ge et al. 2008; Rothacher et al. 2008), is being developed at GFZ for data analysis of various applications, such as real-time deformation monitoring (Network solution mode) and providing service for applications based on PPP (Precise Point Positioning, PPP mode). Figure 12.4 shows the system structure of EPOS-RT. There are three main parts: data communication, processing kernel, and product service. Processing kernel is the main processing unit where observation modelling, parameterization and estimation are encoded. Data communication deals with input and output data. Service part manages and broadcasts products from processing kernel to users.

Various tests and investigations were carried out. In PPP mode, satellite clocks are estimated first, where a comparable precision can be achieved compared to IGS final clocks.

Using the estimated satellite clock and corresponding satellite orbits, ground surface deformation monitoring was carried out. During the Mw7.8 Chile earthquake, on DoY 318, 2007, we analyzed GPS observations at station TALA, 100 km away from the epicenter. Figure 12.5 shows the kinematic PPP results in east component, which is the main deformation at this station. The top plot shows the time series for the whole day, where we see a sudden offset at 15:42 UTC, 1 min after the earthquake hit Tocopilla town. The bottom plot shows the deformations during 20 min before and after the earthquake epoch.

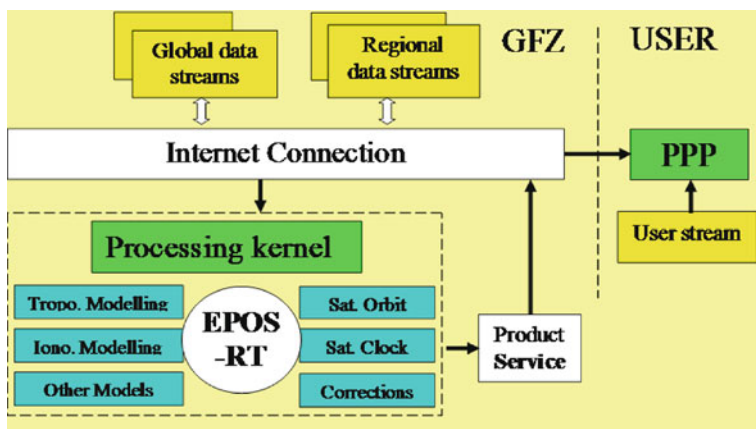


Fig. 12.4 System structure of EPOS-RT

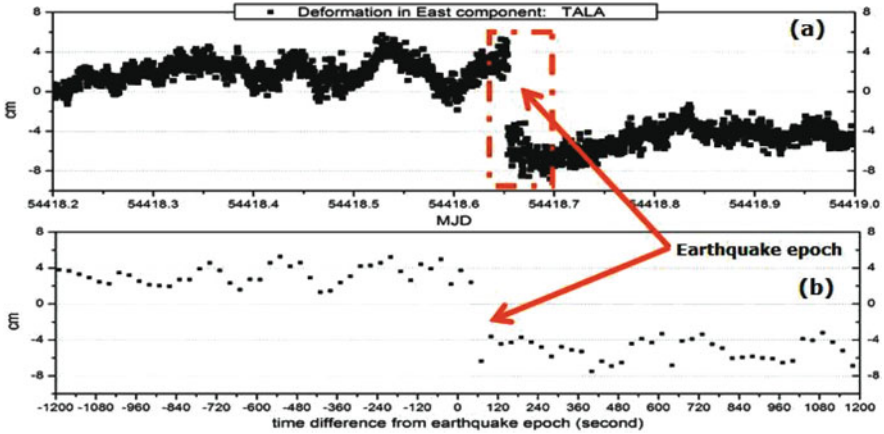


Fig. 12.5 Coordinate changes of TALA during Chile earthquake 2007. (a) Daily PPP solution, displacement of ~8 cm at East component is observed. (b) Coordinate changes of East component during the period 20 min before and after the earthquake epoch

Network solution mode is the second solution mode of EPOS-RT. Making use of the data from the European real-time GNSS network (baselines range from 100 km to 420 km), Fig. 12.6 shows real-time monitoring of the station BZRG, in Bolzano, Italy over 2 days. The system initializes within around 20 min and ambiguity fixing starts thereafter. The real-time coordinate precision is better than 1.3 cm in horizontal components and around 4 cm in height.

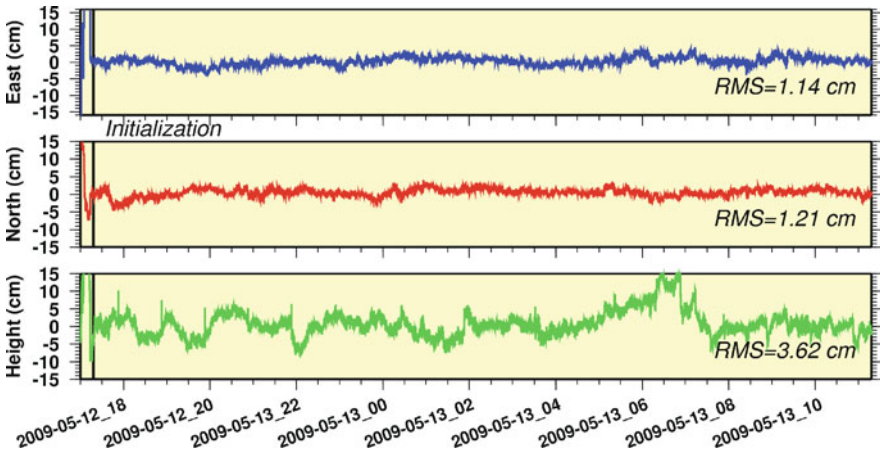


Fig. 12.6 Coordinates differences compared to IGS weekly combination of station BZRG in Italy, where RMS refers to IGS coordinates

12.6 GNSS Atmospheric Sounding

12.6.1 Ground-Based GNSS Meteorology

Crossing the atmosphere, GNSS signals experience a propagation delay depending on constitution of the ionosphere (electron density) and neutral atmosphere (pressure, temperature, water vapor). The neutral atmosphere related zenith path delay (ZPD) above each GPS ground station is a standard product of routine GPS data processing and the humidity induced part of ZPD provides a valuable source of vertically integrated water vapor (IWV) information. Regarding the key role that water vapor plays in the Earth's atmosphere system and the high temporal and spatial variability of water vapor, GPS IWV observations are important for both numerical weather prediction (e.g. Gendt et al. 2004) and climatological investigations (e.g. Nilsson and Elgered 2008). The GPS-based IWV observation technique is characterized by several advantages in comparison to the traditional observing systems: independence on sensor calibrations and therefore long-term stability, all-weather capability, high accuracy and low cost. Based on its global network observations (currently more than 300 stations), the IGS provides ZPD data starting from February 1997. To convert ZPD to IWV meteorological information are normally required, but currently only a limited number of IGS stations are equipped with meteorological sensors. To solve this problem, an analysis technique based on ECMWF (European Centre for Medium-Range Weather Forecasts) analysis data has been developed at GFZ (Heise et al. 2009). Figure 12.7 gives a general view on IGS IWV results derived at GFZ. Good agreement for most of the stations is seen compared to ECMWF.

The globally available GPS data sets are refined by regional observations, e.g. for Europe (<http://egvap.dmi.dk/>) or Germany. The GFZ operates a near real-time analysis center processing about 300 German GPS stations in addition to the IGS

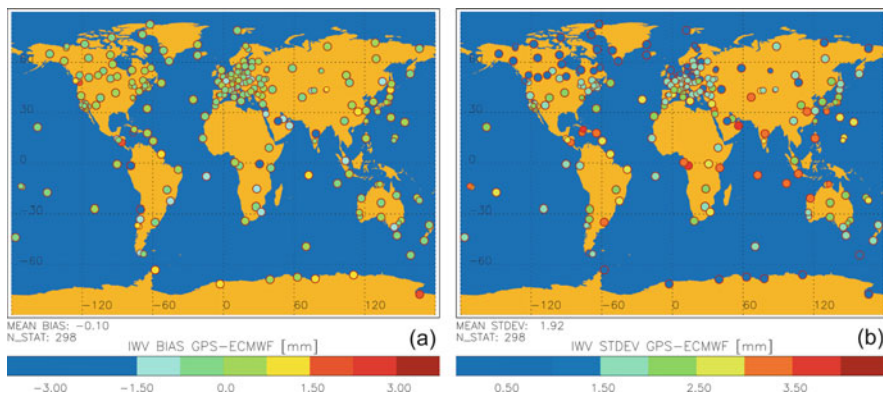


Fig. 12.7 GPS IWV results from 2007 in comparison with ECMWF: (a) bias, (b) standard deviation

and EUREF networks. The EPOS GPS processing package is used to analyze ZPD, IWV and slant delay data at hourly intervals. EPOS is based on a least-squares adjustment of zero-differenced phase and range observations and can either run in PPP or network mode. The Saastamoinen model and the Global or Niell mapping functions are used to apply tropospheric corrections. The un-modelled part of the slant path delay is adjusted for each individual observation to consider local atmospheric inhomogenities. To estimate the IWV, additional meteorological observations are required. The IWV data are available with a temporal resolution of 15 min (Fig. 12.8) and give detailed information about the water vapor distribution above Germany. Furthermore, the slant delays along the satellite-station links are analyzed with a temporal resolution of 2.5 min (Bender et al. 2008). These data provide valuable information about the spatial water vapor distribution. 3D water vapor fields are reconstructed either by using the GPS water vapor tomography (Troller et al. 2006) or by assimilating the slant data to a numerical weather model (Zus et al. 2008).

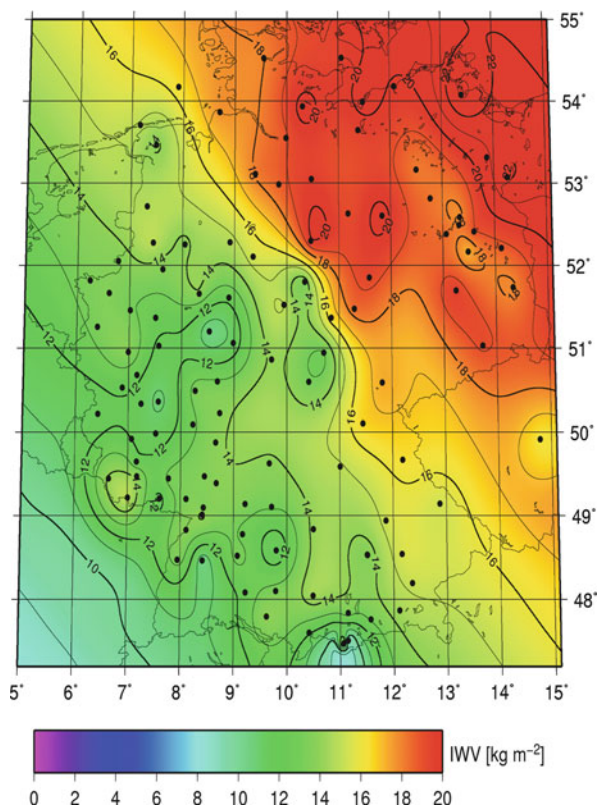


Fig. 12.8 Near real-time IWV distribution above Germany at 1 May 2008, 0:00 UTC. The stations providing data at that time are marked with *black circles*

12.6.2 Spaced-Based Atmosphere Sounding

The availability of GPS radio signals has introduced a new promising remote sensing technique for the Earth's atmosphere. The GPS based radio occultation (RO) exploits these signals received onboard a Low Earth Orbiting (LEO) satellite for atmospheric limb sounding. The GPS signals are influenced by the atmospheric refractivity field resulting in a time delay and path bending of the signal. The atmospheric excess phase is the basic observable that is measured with millimetric accuracy. This is the basis for precise refractivity and temperature profiles (Wickert et al. 2007, 2009; Wickert and Jakowski 2007).

The tropopause layer is one of the key regions of the atmosphere with links to the stratosphere-troposphere exchange as well as climate research. Global mean tropopause height shows an increase in re-analyses and radiosonde observations during the last decades. Tropopause height changes are caused by different forcing mechanisms. One mechanism leading to an increase of the tropopause height is a warming of the troposphere and a cooling of the lower stratosphere. Thus, the tropopause height could be considered as a parameter for the detection of climate change processes and therefore the continuous identification and monitoring of the tropopause height is an important goal in climate research.

The most important data source for the determination of tropopause parameters are radiosonde data whereas model analyses suffer from lower vertical resolution. Despite good vertical resolution of radiosonde measurements a global coverage is impossible. GPS RO enables precise refractivity and temperature profiles with high vertical resolution (< 1 km in the tropopause region). The GPS RO technique requires no active calibration, is weather independent, and the occultations are almost uniformly distributed over the globe. Another important characteristic is the long-term stability of the system, e.g., the CHAMP RO experiment provides data continuously since mid-2001. For the determination of the tropopause different definitions and concepts exist. Here the classical definition of the World Meteorological Organization (WMO) for the first lapse rate tropopause (LRT) derived from a temperature profile is used.

Figure 12.9a shows zonal mean first (dotted) and last (solid) LRT heights derived from CHAMP for the period 2001–2007 for different seasons. The tropopause has a strong meridional structure (LRT1). In the tropics (30°S – 30°N) the tropopause height is nearly constant. In the deep tropics (10°S – 10°N) LRT1 reaches mean values of about 16.5 km. The strongest gradients in the tropopause height occur between 30° and 60° on both hemispheres with mean heights decreasing to 8–10 km (Schmidt et al. 2005). Usually a second tropopause (LRT2) is observed in the extra-tropics during the winter months (Schmidt et al. 2006).

For the first LRT from the CHAMP data between May 2001 and December 2007 a trend analysis was performed showing a global trend of about $+6.6$ m/year (Fig. 12.9b). This value is in excellent agreement with trend results derived from longer radiosonde data sets (Schmidt et al. 2008).

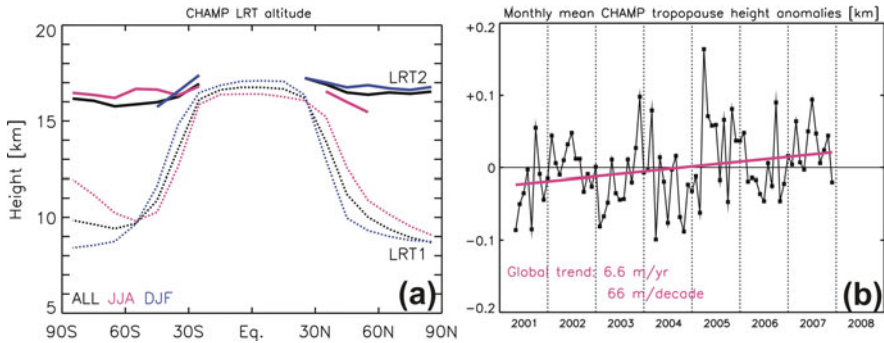


Fig. 12.9 (a) Zonal mean first (*dotted*) and last (*solid*) LRT heights derived from CHAMP for the period 2001–2007 for different seasons (June–August, JJA; December–February, DJF; and global). (b) Monthly global mean CHAMP tropopause height anomalies (2001–2007) and the according linear trend (6.6 m/year)

12.7 Summary

We briefly introduced several GFZ activities related to GNSS based hazard and climate monitoring. These activities are related to ground and satellite based atmospheric sounding, reflectometry, GNSS seismology and corresponding hardware developments. The introduced results underline the ability of GPS, in future extended by Galileo, to be used as a powerful tool for remote sensing to detect natural disasters and climate change related information of the systems Earth.

References

- Bender M, Dick G, Wickert J, Schmidt T, Song S, Gendt G, Ge M, Rothacher M (2008) Validation of GPS slant delays using water vapour radiometers and weather models. *Meteorol Z* 17:807–812
- Beyerle G, Hocke K (2001) Observation and simulation of direct and reflected GPS signals in radio occultation experiments. *Geophys Res Lett* 28:1895–1898
- Chen J, Ge M, Vennebusch M, Deng Z, Gendt G, Rothacher M (2008) Progress of the real-time GNSS software development at GFZ. IGS Workshop 2008, Miami Beach, USA, 2–6 June 2008
- Chen J, Ge M, Dousa J, Gendt G (2009) Evaluation of EPOS-RT for real-time deformation monitoring. *J Global Positioning Syst* 8(1):1–5
- Falck C, Ramatschi M, Subarya C, Bartsch M, Merx A, Hoeberechts J, Rothacher M (2008) GPS Technologies: a new component for tsunami early warning systems. International Conference on Tsunami Warning, Bali, 12–14 November 2008
- Ge M, Chen J, Vennebusch M, Gendt G, Rothacher M (2008) GFZ prototype for GPS-based real-time deformation monitoring. EGU2008, Vienna, Austria, 13–18 April 2008
- Gendt G, Dick G, Reigber C, Tomassini M, Liu Y, Ramatschi M (2004) Near real time GPS water vapor monitoring for numerical weather prediction in Germany. *J Meteorol Soc Jpn* 82:361–370
- GeoForschungsZentrum Potsdam (2006) GPS – SurfacE Deformations withIn Seconds (G-SEIS) Proposal. Approved by BMBF.

- Heise S, Dick G, Gendt G, Schmidt T, Wickert J (2009) Integrated water vapor from IGS ground-based GPS observations: Initial results from a global 5-minute data set. *Ann Geophys* 27(7):2851–2859
- Helm A (2008) Ground-based GPS altimetry with the L1 OpenGPS receiver using carrier phase-delay observations of reflected GPS signals. GFZ German Research Centre for Geosciences, Scientific Technical Report, 8(10)
- Helm A, Stosius R, Beyerle G, Montenbruck O, Rothacher M (2008) Status of GNSS reflectometry related receiver developments and feasibility studies within the German Indonesian Tsunami Early Warning System. IGARSS 2007, Barcelona, Spain, 23–28 July, 2007
- Martín-Neira M (1993) A passive reflectometry and interferometry system (PARIS): application to ocean altimetry. *ESA J* 17:331–355
- Nilsson T, Elgered G (2008) Long-term trends in the atmospheric water vapor content estimated from ground-based GPS data. *J Geophys Res* 113:D19101. doi:10.1029/2008JD010110
- Ramatschi M, Falck C, Subarya C, Bartsch M, Merx A, Hoeberechts J, Rothacher M (2008) Tsunami early warning: benefits from GNSS real-time data transmission. AGU 2008, San Francisco, 15–19 December 2008
- Rothacher M, Babeyko A, Bartsch M, Chen J, Falck C, Galas R, Ge M, Gundlich B, Merx A, Ramatschi M, Schoene T, Vennebusch M (2008) G-SEIS: GPS – SurfacE Deformations with In Seconds, Geotechnologien Statusseminar, Osnabrueck, 8–9 October 2008
- Schmidt T, Heise S, Wickert J, Beyerle G, Reigber C (2005) GPS radio occultation with CHAMP and SAC-C: global monitoring of thermal tropopause parameters. *Atmosph Chem Phys* 5:1473–1488
- Schmidt T, Beyerle G, Heise S, Wickert J, Rothacher M (2006) A climatology of multiple tropopauses derived from GPS radio occultations with CHAMP SAC-C. *Geophys Res Lett* 33. doi:10.1029/2005JD024600
- Schmidt T, Wickert J, Beyerle G, Heise S (2008) Global tropopause height trends estimated from GPS radio occultation data. *Geophys Res Lett* 35. doi:10.1029/2008JD034012
- Stosius R, Helm A, Beyerle G, Höchner A, Wickert J, Rothacher M (2008) Simulation of GNSS-R tsunami detection from space applied to the 2004 Sumatra tsunami event, GNSS-R'08 Workshop on GNSS Reflectometry, ESA/ESTEC, Noordwijk
- Troller M, Geiger A, Brockmann E, Bettems J-M, Bürki B, Kahle H-G (2006) Tomographic determination of the spatial distribution of water vapor using GPS observations. *Adv Spa Res* 37:2211–2217
- Wickert J, Jakowski N (2007) GNSS based sounding of the atmosphere/ionosphere. In: National Report of the Federal Republic of Germany on the Geodetic Activities in the years 2003–2007, ISBN 3769685954, pp 125–129
- Wickert J, Gendt G, Beyerle G, Dick G, Heise S, Helm A, Reigber C, Schmidt T, Jakowski N, Jacobi Ch (2007) Ground and space based GPS atmospheric sounding: brief overview and examples. NARL publication. In: Proceedings of the INTAR Workshop Tirupati, India 2005, pp 1–9, December 2007
- Wickert J, Michalak G, Schmidt T, Beyerle G, Cheng CZ, Healy SB, Heise S, Huang CY, Jakowski N, Köhler W, Mayer C, Offiler D, Ozawa E, Pavelyev AG, Rothacher M, Tapley B, Arras C (2009) GPS radio occultation: results from CHAMP, GRACE and FORMOSAT-3/COSMIC. *Terr Atmos Ocean Sci* 20(1):35–50. doi:10.3319/TAO.2007.12.26.01(F3C)
- Xinhua News Agency (2009) <http://www.xinhuanet.com/xhwenchuan/index.htm>. Accessed on 15 February, 2009
- Zus F, Grzeschik M, Bauer H-S, Wulfmeyer V, Dick G, Bender M (2008) Development and optimization of the IPM MM5 GPS slant path 4DVAR system. *Meteorol Z* 17:867–885

Chapter 13

Satellite Imagery for Landslide Mapping in an Earthquake-Struck Area

Xiaojun Yang

Abstract This study examines the utilities of satellite imagery for landslide mapping in connection to a recent major earthquake event. The study area covers part of the Wenchuan County, Sichuan Province, China, approximately 30 km away from the epicenter of the Sichuan Earthquake that struck on 12 May 2008. Estimated at the magnitude of 7.9 on the Richter scale, this earthquake is cited as the 19th deadliest earthquake of all time. The Sichuan Earthquake and its aftershocks have triggered numerous landslides that were directly responsible for at least one-third of the overall casualties and widespread infrastructure damage. The primary data used are two satellite images acquired before and after the Sichuan Earthquake by Terra's Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Landsat 5 Thematic Mapper (TM). Georeferencing and radiometric normalization are conducted before further processing the satellite scenes. A Normalized Difference Vegetation Index (NDVI) image is computed from each of the images, and the NDVI change after the Earthquake is analyzed. It is found that the areas with a large NDVI decrease were largely related to earthquake-triggered landslide activities. This study demonstrates that satellite imagery can be quite useful to map the spatial distribution of earthquake-induced landslides quickly.

13.1 Introduction

Landslides are rock, earth, or debris flows on hillslopes due to gravity. Globally, they constitute a major type of geological hazards that are widespread, claiming multiple billions of dollars in infrastructure and property damages and numerous fatalities every year (Dai et al. 2002; Fourniadis et al. 2007; Huang 2009; Singhroy and Molch 2004). Expansion of urban and recreational development into hillslope areas lead to more people that are threatened by landslide hazards (Alexander 1989; Oliver 1993; Switalski et al. 2004; Tarantino et al. 2007). Landslides are often triggered by other

X. Yang (✉)

Department of Geography, Florida State University, Tallahassee, FL 32306, USA

e-mail: xyang@fsu.edu

major natural disasters or extreme events, such as earthquakes, volcanoes, flooding, and tsunamis (Keefer 2002; Sato and Harp 2009; Wasowski 1998). Monitoring landslide activities and predicting spatio-temporal slope failures can help understand the nature of terrain failure, formulate mitigation strategies, and ultimately minimize the losses from landslide hazards (Havenith et al. 2006; Mantovani et al. 1996; Van Westen et al. 2008).

Concerned with the spatial distribution of slope failures, landslide mapping is essential for slope instability investigations (Alcantara-Ayala et al. 2006; Domakinis et al. 2008; Wang et al. 2009). Both ground surveys and remote sensing can be used for landslide mapping (Brardinoni et al. 2003). Ground surveys can help obtain detailed, accurate information on landslide distribution and classification but are limited by their local scales and logistical constraints when investigating a large area or an area affected by one or more extreme events. Remote sensing, through the use of cameras and sensors mounted on aerospace-borne platforms, can help overcome these limitations.

With excellent spatial resolution and stereoscopic viewing, aerial photographs have extensively been used in landslide mapping and zonation (e.g. Brardinoni et al. 2003; Chen et al. 2006; Fookes et al. 1991; Harp and Jibson 1996; Hearn 1995; Whitworth et al. 2005). Recent innovations in remote sensing have permitted geoscientists with new data to advance landslide studies. Examples of such data include high-resolution satellite imagery (e.g. Hervas et al. 2003; Nichol et al. 2006), radar imagery (e.g. Colesanti and Wasowski 2006; Singhroy and Molch 2004), and lidar point clouds (e.g. Chen et al. 2006; Van Den Eeckhaut et al. 2007). These recently available remote sensor data are with high spatial resolutions, particularly suitable for landslide identification and classification (Tsutsui et al. 2007). Because of the on-demand acquisition, however, they may not be available for some areas where numerous landslides occur due to extreme events such as earthquakes. On the other hand, such high resolutions may not need for reconnaissance purposes (Weirich and Blesius 2007).

Archival satellite imagery series are often collected in a systematic mode, allowing a retrospective analysis of landslide activities (Roessner et al. 2005). Their medium resolution and synoptic coverage allow a regional-scale assessment of landslide distribution (Weirich and Blesius 2007). Moreover, the parallel development in image processing and information extraction has substantially added the values of satellite imagery in landslide studies on different scales (e.g. Borghuis et al. 2007; Chang et al. 2007; Marcelino et al. 2009; Nichol and Wong 2005). While the potential of satellite imagery has been recognized, further research will certainly be maintained and will probably intensify in order to adopt these techniques for landslide mapping in a productive mode, thus reinforcing the absolute and comparative utility of modern remote sensing technology.

The objective of this study was to identify a method for mapping landslide distribution in connection to a recent major earthquake event. This method was mainly based on the use of archival satellite imagery and image processing techniques. The following sections will provide an overview on the case study site, detail the research methodology, and analyze the results.

13.2 Study Area

The case study site covers part of the Wenchuan County, Sichuan Province, China, approximately 30 km from the epicenter of the Sichuan Earthquake that struck at 14:28:01.42 local time (06:28:01.42 Coordinated Universal Time-UTC) on 12 May 2008, estimated at the magnitude of 7.9 on the Richter scale (USGS 2008). Cited as China's most devastating earthquake in more than three decades, the Sichuan Earthquake killed 69227, injured 374,643, and made 4.8 million people homeless. Figure 13.1 shows the location of the study site, along with the epicenter.

Physiographically, the study site is situated at the eastern edge of the Tibetan Plateau. It is located at the Longmenshan fault zone, a northeast striking thrust structure defining the boundary between the high topography of the Tibetan Plateau to the west and the relatively undeformed Sichuan Basin to the east (Hubbard and Shaw 2009). The Sichuan Earthquake of 12 May 2008 occurred as the result of the motion of the Longmenshan fault, predominately on its mid fracture known as the Yingxiu-Beichuan fracture. The displacement of this fault zone was modeled at a

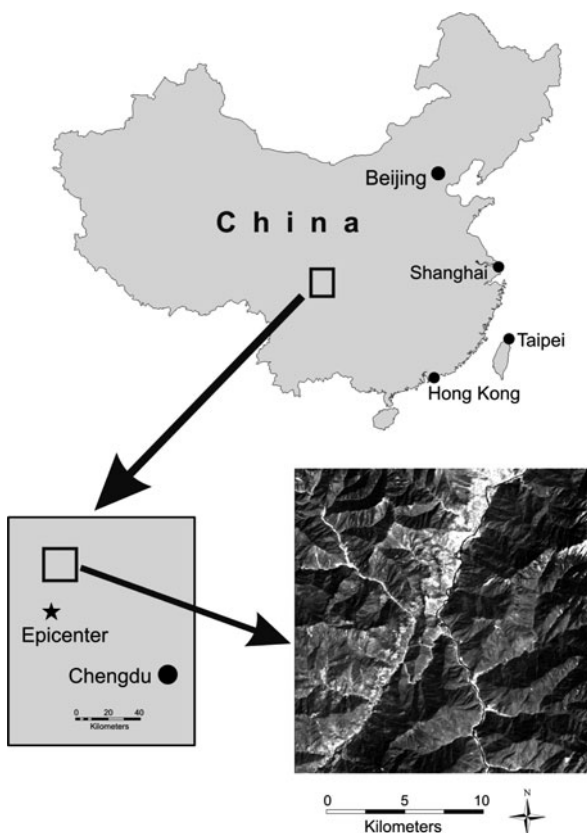


Fig. 13.1 Location of the study area. The earthquake epicenter shown (*lower left*). The lower right figure is a subset of Band 7 of the Landsat TM scene acquired on 18 September 2007, covering the entire study area



Fig. 13.2 Selected photos showing damages after the earthquake struck China's Sichuan Province. With more than 69227 casualties, this quake is cited as the worst natural disaster striking China since the Tangshan Earthquake in 1976. Photos were taken near the epicenter. Photographed by Xiaojun Yang



Fig. 13.3 Landslide development after the Sichuan Earthquake. Photographs were taken from Doujiangyan to Yingxiu, which are close to the epicenter. Numerous landslides have been triggered mostly along streams and roads. Photographed by Xiaojun Yang

maximum of 9 m, which generated deformations of the surface greater than 3 m and increased the stress at the northeastern and southwestern ends of the fault (USGS 2008).

With the focus of 19 km in depth, the Sichuan Earthquake and its aftershocks created a rupture zone extending more than 200 km. They have triggered more than 15000 incidences of rockfalls, debris flows, and other types of landslides, along with more than 10000 potential rockfall sites (Yin et al. 2009). These landslides have been directly responsible for more than 20000 casualties and widespread infrastructure damage (Figs. 13.2 and 13.3).

13.3 Research Methodology

The research methodology included several major components: data collection, image preprocessing, image transformation, change detection, thematic accuracy assessment, and the production of final landslide distribution map (Fig. 13.4).

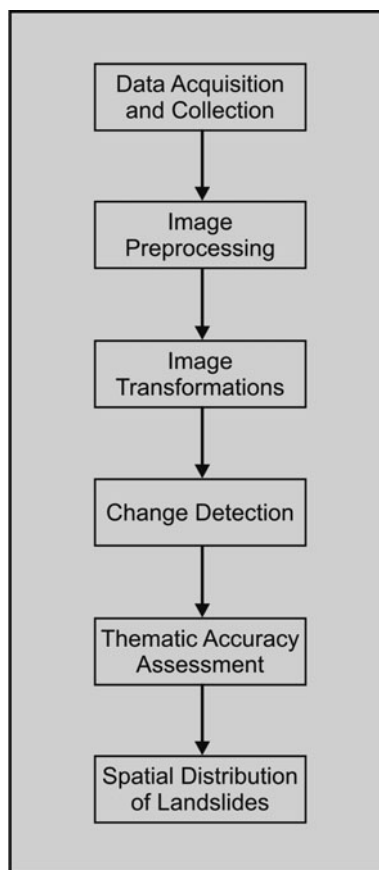


Fig. 13.4 Flowchart of the working procedural route adopted in this study. Note that only major procedures are illustrated here

13.3.1 Data Acquisition and Collection

Ideally, images from a single remote sensor are most desired for landscape change analysis. But this may not always be possible due to other environmental and technological constraints. For example, it is always difficult to obtain cloud-free satellite scenes over mountainous areas due to high moistures, particularly after an extreme event such as earthquakes or severe storms. After searching the entire Landsat database archived by USGS EROS Data Center, we were able to obtain a cloud-free Thematic Mapper (TM) scene acquired at 3:32:14 UTC on 18 September 2007, slightly more than eight months before the Sichuan Earthquake. We could not find any post-event TM scenes with less cloud cover for our study site. Note that we excluded the images from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) because of the scan line corrector failure since 31 May 2003. Fortunately, we were able to find a good post-event scene acquired by Terra's Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). The ASTER scene was acquired at 3:57:00 UTC on 23 May 2008, 11 days after the Sichuan Earthquake. The TM and ASTER scenes were used as the primary data in this study. The specific dates, sensors, satellite scene ID or reference system, resolutions, and other environmental parameters are summarized in Table 13.1.

Table 13.1 Characteristics of the satellite images used

Data provider	USGS EROS Data center		USGS and Japan ASTER Program	
Satellite	Landsat 5		Terra	
Sensor	Thematic Mapper (TM)		Advanced Spaceborne thermal Emission and Radio- meter (ASTER)	
Scene ID or reference system	Path 130 Row 38		035700	035709
Acquisition Time	18 September 2007 3:32:14 UTC		23 May 2008 3:57:00 UTC	23 May 2008 3:57:09 UTC
Bands	VNIR (1-4 bands) SWIR (bands 5 and 7) TIR (band 6)		VNIR (1-3 bands) SWIR (4-9 bands) ^a TIR (10-14 bands)	
Pixel size	28.5 m (120 m for band 6)		VNIR: 15 m; SWIR: 30 m; TIR: 60 m	
Radiometric resolution	8-bits		VNIR and SWIR: 8-bits; TIR: 12-bits	
Swath width	185 km		60 km	
Solar azimuth	140.34°		122.274°	120.878°
Solar elevation	53.84°		71.868°	72.028°

^aThe six ASTER SWIR bands are not useable because of saturation of values and severe striping. This is due to the malfunctioning of the sensors as a result of anomalously high SWIR detector temperatures since May 2008 (http://www.science.aster.ersdac.or.jp/en/about_aster/swir_en.pdf)

In addition to the above satellite images, we collected diverse geospatial datasets, including the epicenter location, administration boundaries, digital elevation model (DEM) data derived from the Shuttle Radar Topography Mission (SRTM), geological maps, socio-economic data, and so on. These geographically referenced data were used to facilitate satellite image-based landslide mapping at different stages.

We also conducted a limited field work to collect first-hand data concerning the earthquake impacts and landslide development in Wenchuan. This *in situ* investigation has been quite helpful for processing and interpretation of the satellite images.

13.3.2 Image Preprocessing

Georeferencing, mosaicking, subset, and radiometric normalization were carried out in the phase of image preprocessing. Establishing a common georeferencing system for all data layers is a prerequisite for spatially corrected landscape change mapping. The two dates of image scenes were geometrically rectified by the data providers, namely, USGS EROS Data Center and NASA. Here, each scene was further georeferenced to the Universal Transverse Mercator (UTM) map projection (Zone 48 N), the World Geodetic System (WGS) 1984 horizontal datum, and the WGS 1984 ellipsoid. Other geospatial data layers were also georeferenced similarly so that further integration and analysis can be possible.

A mosaicking procedure was used to combine the two adjacent ASTER scenes in order to cover the entire study area (see Fig. 13.1). The actual image subset used for landslide mapping covers a rectangle area of 18586 m in width and 19187 m in height (Fig. 13.5). Therefore, both the TM and mosaicked ASTER images were clipped with the rectangle as the mask. Note our image subset operation also

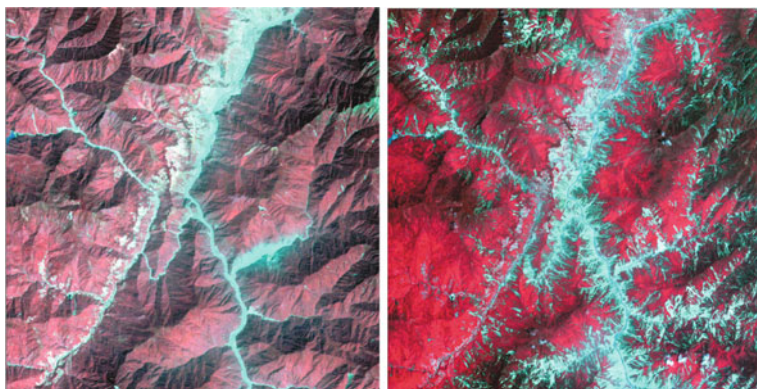


Fig. 13.5 Remote sensor data used in this study. The left is part of a Landsat TM scene acquired on 18 September 2007, displayed in a false color composite with bands 1, 2 and 4; the right is part of an ASTER image obtained on 23 May 2008, in a false color display with bands 1–3. Each image is north-south oriented. The dimension of each image is 18586 m in width and 19187 m in height

included the exclusion of the thermal bands from the TM and ASTER images for further analysis due to their coarse spatial resolution. For the ASTER scenes, the six SWIR bands (see Table 13.1) were also excluded for further analysis because of their poor image quality caused by the malfunctioning of the SWIR detectors since 2007.

The two images were acquired by different sensors at different dates. To allow meaningful detection of landscape changes based on these images, a common radiometric response between them should be restored (Yang and Lo 2000). For this purpose, the relative radiometric normalization (RRN) is preferred over the absolute radiometric correction method because no in situ atmospheric data at the time of satellite overpasses are necessary. Based on the comparative research done by Yang and Lo (2000), the RRN procedure proposed by Hall et al. (1991) was applied to the two images in order to suppress their radiometric differences caused by the variations among atmospheric conditions, sensor-target-viewing geometry, vegetation growing seasons, and phenological characteristics. With the TM image dated on 18 September 2007 as the reference, the ASTER image dated on 23 May 2008 as the subject scene was then radiometrically rectified by using radiometric control sets (see Hall et al. 1991).

13.3.3 Image Transformation

In this study, the image-to-image comparison approach was used to detect the spatial occurrences of earthquake-induced landslides. This pre-classification change detection method involves the use of either original images or derived images in a ratioing or differencing algorithm. The use of original image bands seems to be quite straightforward but has two major drawbacks. Firstly, when using multispectral imagery for change detection, it may not be efficient to perform a band-to-band comparison exhaustively for all bands between two dates. It is more desired to use only a smaller number of bands for change detection. However, there is no explicit guidance that could be used to choose specific bands in this regard. Secondly, change detection using image bands can be complicated by the variation of image signals caused by other external influences such as atmospheric or sun angle differences. This can be true even when radiometrically normalized images are used. Therefore, it is difficult to interpret the detected changes. To overcome these drawbacks, we used derived images in change detection.

There are some different ways to transform multispectral images into a smaller and easier to interpret set that represents most of the information from the original dataset. Some popular image transformation techniques include principal component analysis (PCA), independent component analysis (ICA), minimum noise fraction (MNF), band ratioing, and a large number of vegetation indices (Jensen 2005). Our major interest was to examine the vegetation changes before and after the Earthquake, which were further used to pinpoint the spatial distribution of landslide activities. Therefore, the use of a popular vegetation index became a natural choice. To this end, we computed Normalized Difference Vegetation Index (NDVI) from each of the TM and radiometrically normalized ASTER scenes. NDVI has

widely been used to quantify vegetation characteristics from remote sensor data. It was computed by using the following equation:

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}},$$

where RED and NIR represent the spectral reflectance measurements obtained in the red and near-infrared regions, respectively. RED and NIR correspond to Bands 3 and 4 for the TM scene and Bands 2 and 3 for the ASTER scene. Note either image digital numbers or reflectance can be used for the above computation. In any cases, the resultant NDVI image ranges from -1.0 to 1.0 in its values.

13.3.4 Change Detection

An image differencing procedure was implemented to detect the NDVI change by using the following equation:

$$\Delta\text{NDVI} = \text{NDVI}_2 - \text{NDVI}_1,$$

where ΔNDVI is the change in NDVI value, NDVI_2 is the NDVI image derived from the radiometrically normalized ASTER scene acquired 11 days after the Sichuan Earthquake, and NDVI_1 is the NDVI image from the TM scene obtained slightly more than 8 months before the catastrophic event. The resulting NDVI change image was used to create an initial landslide distribution map by highlighting the pixels with a decrease of at least 10% in their brightness values as a direct result of the subtraction. The determination of 10% as the threshold was made through an interactive process using a visual image interpretation method. Note that setting a threshold to highlight changes, as a standard practice in an image-differencing change detection analysis (Jensen 2005), has also been used to map landslides (e.g. Hervas et al. 2003; Rosin et al. 2000). In addition to the interactive method, some more robust yet complicated methods such as the ones discussed by Rosin and Ellis (1995) could be used to define the threshold value.

One more step of processing was carried out to suppress the boundary errors that occur at the landslide distribution boundaries due to the occurrence of spectral mixing within a pixel, image noises, and geometric registration errors of the two input images (Yang and Lo 2002). These boundary errors are often small and in the form of salt and pepper. These small areas have to be removed and replaced with class values based on their surroundings. A modified 3 by 3 modal filter was used to reduce boundary errors in the initial landslide distribution map. The modal filter is also called focal majority filter. It is applied to an $n \times n$ pixel patch, where n is an odd integer. A histogram of class values in the patch is generated and the value having the highest frequency is returned as the new central value. The centre pixel's value thus becomes that of the most commonly occurring class within the patch. In

this way, the small (and erroneously classified) pixels are reclassified according to the dominant class within the patch (Yang and Lo 2002).

13.3.5 Thematic Accuracy Assessment

A standard procedure for thematic accuracy assessment recommended by Congalton (1991) was implemented here. A total of 100 test points were chosen by using a stratified random sampling scheme. Each of these points was assessed by using a visual image interpretation procedure based on image elements and the reference data from field surveys and Google Earth. An error matrix was constructed, and standard accuracy report metrics were generated, including producer accuracy (87.04%), user accuracy (94.00%), overall accuracy (90.00%), overall kappa index (0.80), and conditional kappa index (0.87). The user accuracy indicates that approximately 6.0% of landscape changes that were not caused by landslide activities were classified as landslides. The producer accuracy reveals that approximately 13% of landslide-derived landscape changes were not correctly classified as landslides. Overall, most of the landslides have successfully been detected and mapped.

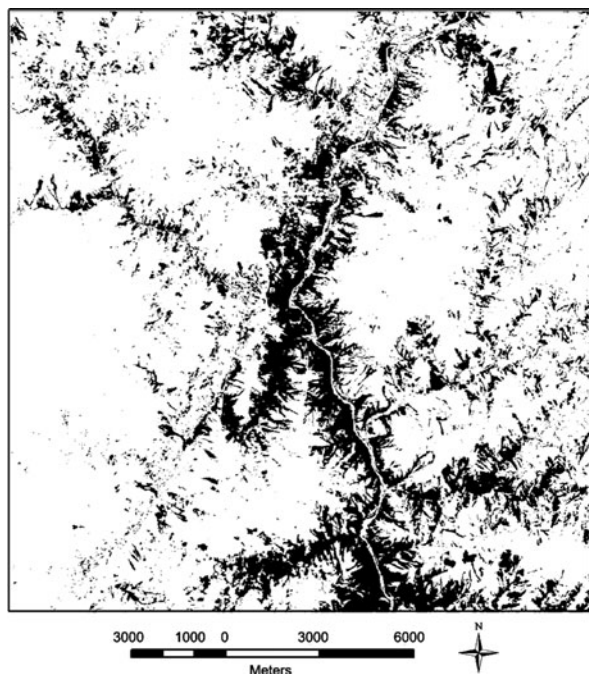


Fig. 13.6 Spatial distribution of the landslides triggered by the earthquake struck on 12 May 2008. Note that most of these landslides (*black patches*) occurred along the Mingjing River and its major tributaries

This is a good indication that the image processing and change detection procedures adopted in our study have been effective in mapping the spatial distribution of earthquake-triggered landslides from satellite imagery.

13.4 Results

Figure 13.6 illustrates the spatial distribution of landslides after the Earthquake. Quantitatively, these earthquake-triggered landslides occupied 5903 hectares or approximately 16.48% of the total study area. By using the original satellite scenes and Google Earth, we visually examined the landslide distribution in relation to other geographic features such as stream networks, terrain slope, and road networks. We found that most of the landslides occurred along the Mingjing River and its major tributaries (black patches in Fig. 13.6). Also some landslides occurred around mountain summits and ridges, indicating the seismic shaking was quite strong during the Sichuan Earthquake and its aftershocks. This distribution pattern was also observed by other investigators who have done extensive field surveys, such as Sato and Harp (2009) and Yin et al. (2009).

13.5 Conclusions

In this study, we have demonstrated how satellite imagery can be used to map the spatial distribution of landslides in connection to a recent major earthquake event. Critical to this study has been the acquisition of the two dates of satellite scenes with one obtained before and the other after the catastrophic event. Ideally, these scenes should be acquired by the same sensors immediately before and after the earthquake event. But this is not always possible due to some environmental and logistical constraints. Radiometric normalization of the satellite imagery acquired by different sensors at different dates has been crucial in the image-differencing change detection procedure adopted in this study. The use of derived images rather than image bands has improved the manageability and interpretability of the detected changes in the context of strong seismic activities. The thematic accuracy assessment indicates that the image processing and change detection procedures identified here have been effective in mapping the distribution of earthquake-triggered landslides. We found that most of the large landslides triggered by the Earthquake occurred along the Mingjiang River and its major tributaries. Some landslides occurred around mountain summits and ridges, suggesting the strong impacts of seismic shaking during the Sichuan Earthquake and its aftershocks. These remote sensing-based observations are consistent with the findings obtained through extensive field surveys.

Given the relatively simple procedures and the good accuracy, the landslide mapping method identified in this study seems to be quite promising from an operational perspective. This method can be useful for rapidly mapping landslide distribution in connection to different catastrophic events, such as earthquakes, extreme storms, or

tsunamis, which tend to occur more frequently from a global perspective. Mapping landslide distribution can further help formulate geohazard mitigation strategies.

Nevertheless, we also note some limitations with the landslide mapping method identified in this study, which were mostly related to the validity of using a threshold value to pinpoint landslide-caused landscape changes. Further research is needed to develop more robust methods that can help define this threshold value efficiently.

Acknowledgments The author would like to thank Florida State University for the time release in conducting this work. The research was partially supported by the Florida State University Council on Research and Creativity and the Chinese Academy of Sciences through the International Partnership Project Ecosystem Processes and Services. Thanks are also due to the Global Land Cover Facilities for sharing their data collection for the 2008 China Earthquake. Lastly, the author wishes to thank the anonymous reviewer and Dr. Emilio Chuvieco for their time and effort in reviewing the earlier version of the manuscript, which helped improve the scholarly quality of this paper.

References

- Alcantara-Ayala I, Esteban-Chavez O, Parrot JF (2006) Landsliding related to land-cover change: A diachronic analysis of hillslope instability distribution in the Sierra Norte, Puebla, Mexico. *Catena* 2:152–165
- Alexander D (1989) Urban landslides. *Prog Phys Geog* 2:157–191
- Borghuis AM, Chang K, Lee HY (2007) Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. *Int J Remote Sens* 7–8:1843–1856
- Brardinoni F, Slaymakerl O, Hassan MA (2003) Landslide inventory in a rugged forested watershed: A comparison between air-photo and field survey data. *Geomorphology* 3–4:179–196
- Chang YL, Liang LS, Han CC, Fang JP, Liang WY, Chen KS (2007) Multisource data fusion for landslide classification using generalized positive Boolean functions. *IEEE Trans Geosci Remote* 6:1697–1708
- Chen RF, Chang KJ, Angelier J, Chan YC, Deffontaines B, Lee CT, Lin ML (2006) Topographical changes revealed by high-resolution airborne LiDAR data: The 1999 Tsaoling landslide induced by the Chi–Chi earthquake. *Eng Geol* 3–4:160–172
- Colesanti C, Wasowski J (2006) Investigating landslides with space-borne synthetic aperture radar (SAR) interferometry. *Eng Geol* 3–4:173–199
- Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens Environ* 1:35–46
- Dai FC, Lee CF, Ngai YY (2002) Landslide risk assessment and management: An overview. *Eng Geol* 1:65–87
- Domakinis C, Oikonomidis D, Astaras T (2008) Landslide mapping in the coastal area between the Strymonic Gulf and Kavala (Macedonia, Greece) with the aid of remote sensing and geographical information systems. *Int J Remote Sens* 23:6893–6915
- Fookes PG, Dale SG, Land JM (1991) Some observations on a comparative aerial-photography interpretation of a landslipped area. *Q J Eng Geol* 3:249–265
- Fourniadis IG, Liu JG, Mason PJ (2007) Landslide hazard assessment in the Three Gorges area, China, using ASTER imagery: Wushan-Badong. *Geomorphology* 1–2:126–144
- Hall FG, Strebel DE, Nickeson JE, Goetz SJ (1991) Radiometric rectification – toward a common radiometric response among multitemporal, multisensor images. *Remote Sens Environ* 1:11–27
- Harp EL, Jibson RW (1996) Landslides triggered by the 1994 Northridge, California, earthquake. *B Seismol Soc Am* 1:S319–S332
- Havenith HB, Torgoev I, Meleshko A, Alioshin Y, Torgoev A, Danneels G (2006) Landslides in the Mailuu-Suu Valley, Kyrgyzstan – Hazards and impacts. *Landslides* 2:137–147

- Hearn GJ (1995) Landslide and erosion hazard mapping at Ok-Tedi copper mine, Papua-New-Guinea. *Q J Eng Geol* 28:47–60
- Hervas J, Barredo JI, Rosin PL, Pasuto A, Mantovani F, Silvano S (2003) Monitoring landslides from optical remotely sensed imagery: The case history of Tessina landslide, Italy. *Geomorphology* 1–2:63–75
- Huang RQ (2009) Some catastrophic landslides since the twentieth century in the southwest of China. *Landslides* 1:69–81.
- Hubbard J, Shaw JH (2009) Uplift of the Longmen Shan and Tibetan plateau, and the 2008 Wenchuan (M=7.9) earthquake. *Nature* 7235:194–197
- Jensen JR (2005). *Introductory digital image processing: a remote sensing perspective*, 3rd edn. Pearson Prentice Hall, NJ
- Keefer DK (2002) Investigating landslides caused by earthquakes – A historical review. *Surv Geophys* 6:473–510
- Mantovani F, Soeters R, Van Westen CJ (1996) Remote sensing techniques for landslide studies and hazard zonation in Europe. *Geomorphology* 3–4:213–225
- Marcelino EV, Formaggio AR, Maeda EE (2009) Landslide inventory using image fusion techniques in Brazil. *Int J Appl Earth Observ Geoinf* 3:181–191
- Nichol J, Wong MS (2005) Satellite remote sensing for detailed landslide inventories using change detection and image fusion. *Int J Remote Sens* 9:1913–1926
- Nichol JE, Shaker A, Wong MS (2006) Application of high-resolution stereo satellite images to detailed landslide hazard assessment. *Geomorphology* 1–2:68–75
- Oliver S (1993) 20th-century urban landslides in the Basilicata region of Italy. *Environ Manage* 4:433–444
- Roessner S, Wetzel HU, Kaufmann H, Sarnagoev A (2005) Potential of satellite remote sensing and GIS for landslide hazard assessment in southern kyrgyzstan (Central Asia). *Nat Hazards* 3:395–416
- Rosin PL, Ellis T (1995) Image difference threshold strategies and shadow detection. In: *Proceedings of the 1995 British conference on Machine vision (Vol. 1)*, BMVA Press Surrey, UK, pp. 347–356
- Rosin PL, Hervás J, Barredo JI (2000) Remote sensing image thresholding for landslide motion detection. In: *Proceedings of the 1st International Workshop on Pattern Recognition Techniques in Remote Sensing*, Andorra, pp. 10–17
- Sato HP, Harp EL (2009) Interpretation of earthquake-induced landslides triggered by the 12 May 2008, M7.9 Wenchuan earthquake in the Beichuan area, Sichuan Province, China using satellite imagery and Google Earth. *Landslides* 2:153–159
- Singhroy V, Molch K (2004) Characterizing and monitoring rockslides from SAR techniques. *Adv Spa Res* 33:290–295
- Switalski TA, Bissonette JA, DeLuca TH, Luce CH, Madej MA (2004) Benefits and impacts of road removal. *Front Ecol Environ* 1:21–28
- Tarantino C, Blonda P, Pasquariello G (2007) Remote sensed data for automatic detection of land-use changes due to human activity in support to landslide studies. *Nat Hazards* 1: 245–267
- Tsutsui K, Rokugawa S, Nakagawa H, Miyazaki S, Cheng CT, Shiraishi T, Yang SD (2007) Detection and volume estimation of large-scale landslides based on elevation-change analysis using DEMs extracted from high-resolution satellite stereo imagery. *IEEE Trans Geosci Remote* 6:1681–1696
- USGS (2008) Magnitude 7.9 – eastern Sichuan, China: 2008 May 12 06:28:01 UTC [Online] (Updated 22 May 2009). <http://earthquake.usgs.gov/eqcenter/eqinthenews/2008/us2008ryan/us2008ryan.php>. Accessed 18 July 2009
- Van Den Eeckhaut M, Poesen J, Verstraeten G, Vanacker V, Nyssen J, Moeyersons J, van Beek LPH, Vandekerckhove L (2007) Use of LIDAR-derived images for mapping old landslides under forest. *Earth Surf Proc Land* 5:754–769

- Van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Eng Geol* 3–4:112–131
- Wang FW, Cheng QG, Highland L, Miyajima M, Wang HB, Yan CG (2009) Preliminary investigation of some large landslides triggered by the 2008 Wenchuan earthquake, Sichuan Province, China. *Landslides* 1:47–54
- Wasowski J (1998) Understanding rainfall-landslide relationships in man-modified environments: A case-history from Caramanico Terme, Italy. *Environ Geol* 2–3:197–209
- Weirich F, Blesius L (2007) Comparison of satellite and air photo based landslide susceptibility maps. *Geomorphology* 4:352–364
- Whitworth MCZ, Giles DP, Murphy W (2005) Airborne remote sensing for landslide hazard assessment: A case study on the Jurassic escarpment slopes of Worcestershire, UK. *Q J Eng Geol Hydroge* 38:285–300
- Yang X, Lo CP (2000) Relative radiometric normalization performance for change detection from multi-date satellite images. *Photogramm Eng Rem S* 8:967–980
- Yang X, Lo CP (2002) Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. *Int J Remote Sens* 9:1775–1798
- Yin YP, Wang FW, Sun P (2009) Landslide hazards triggered by the 2008 Wenchuan earthquake, Sichuan, China. *Landslides* 2:139–152

Chapter 14

Relations Between Human Factors and Global Fire Activity

Emilio Chuvieco and Chris Justice

Abstract Biomass burning is a critical factor to understand both atmospheric and vegetation properties worldwide. Recent changes in global temperatures and socio-economic transformations have affected traditional fire regimes, thus magnifying the negative effects of fire upon human and ecological values. Most sources recognize the importance of human factors in fire ignition, but few studies have tried to understand human patterns of fire at global scale. This paper addresses some of those factors, by using geographical databases covering the whole planet. Fire occurrence was estimated from a database of hot-spots detected by the MODIS sensor, covering the period from 2001 to 2006. Human factors are proven to be related to fire persistency and seasonality, while fire density patterns are associated to human variables for specific climates and vegetation covers.

14.1 Introduction

Most international reports on biomass burning recognize the importance of the human factors in fire occurrence (FAO 2007). Although fire is a natural factor in many ecosystems, human activities play a critical role in altering natural fire conditions, either by increasing ignitions (Leone et al. 2003), or by suppressing natural fires (Johnson et al. 2001; Keeley et al. 1999). Both factors are contradictory, and act mainly through the mixture of fire policy practices, on one hand, and land uses and demographic changes on the other.

Most develop countries have maintained for several decades a fire suppression policy, which has lead to almost total fire exclusion. The long term impact of that policy has implied an alteration of traditional fire regimes, commonly by increasing average burn severity and size, as a result of higher fuel accumulation. Pyne (2001), although other authors are more critical about the real implication of fire

E. Chuvieco (✉)

Department of Geography, University of Alcalá, Colegios 2, 28801 Alcalá de Henares, Spain
e-mail: emilio.chuvieco@uah.es

suppression policy (Johnson et al. 2001), or they tend to put more emphasis on the impact of climate changes (Westerling et al. 2006). Parallel to that trend, the new urbanization patterns in developed countries tend to favor the location of houses in or close to forested areas, and therefore increasing both the ignition probability of those areas (as a result of carelessness) and their vulnerability, as values at stake increase (Cortner et al. 1990; Pyne 2001).

For developing countries, fire is still the most common tool for land clearing, and therefore it is strongly associated to deforestation, especially in Tropical areas (Cochrane et al. 1999; DeFries et al. 2002). The traditional use of fire in shifting cultivation has turned in the last decades to permanent land use change, in favor of cropland and grasslands. In addition, fire is a traditional tool to manage permanent grasslands, which are burned annually to favor new shoots and improve palatability (Hobbs et al. 1991).

Global and local implications of changing natural fire conditions have been widely recognized, with major effects on air quality, greenhouse gas emissions, soil degradation and vegetation succession (Goetz et al. 2006; Parisien et al. 2006; Randerson et al. 2005). The role of human activities in changing those conditions has not been assessed at global scale. Several local studies have identified factors that are commonly associated to human fire ignition, such as distance to roads, forest-agricultural or forest-urban interfaces, land use management, and social conflicts (unemployment, rural poverty, hunting disputes,) (Leone et al. 2003; Martínez et al. 2009; Vega-García et al. 1995). On the other hand, humans not only cause fires, but they suffer their consequences as well. Fire is recognized as a major natural hazard (FAO 2007), which imply severe losses of human lives, properties and other socio-economic values (Radeloff et al. 2005; Reisen and Brown 2006).

The purpose of this paper is to analyze the relations between global fire activity and human variables that both affect the causes and the impacts of biomass burning. Since the official fire statistics provided by national bodies are mostly very inaccurate or non-existent (FAO 2007), the only efficient tool to assess fire activity at global scale is the use of hot-spots detected from satellite data. A 6-year series of Terra-MODIS active fire detections have been used to compute several fire metrics, which have been correlated to different human variables to identify global trends between human and fire activity. Previous studies based also on satellite data have identified the impact of agriculture in spatial patterns of fire activity in Latin America (Di Bella et al. 2006), as well as the relation between fire occurrence and distances to roads and populated areas in Russia (Mollicone et al. 2006).

14.2 Methods

14.2.1 Active Fire Database

For this study, 6 years of Terra-MODIS data (2001–2006) were used as data input to characterize global fire activity. The following metrics were used (Chuvieco et al. 2008):

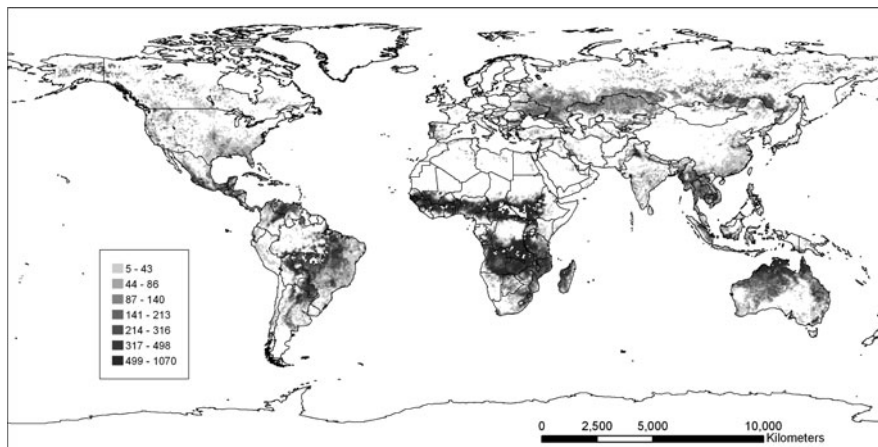


Fig. 14.1 Average monthly density of fires detected from the MODIS sensor (in fire counts 10^{-5} km^2). Data have been grouped in 0.5×0.5 degrees cells (after Chuvieco et al. 2008). To avoid working with very small fraction numbers, the original density was multiplied by 10^5 to render integer numbers in the range between 4 and 3862 (meaning a fire density between 0.00004 and 0.03862 fire counts km^{-2} month^{-1})

1. Average Fire density (AFD, counts km^{-2} month^{-1}), defined as the mean density of detected fires over all years and months (Fig. 14.1). It was derived as an indicator of fire density.
2. Length of Fire Period (LFP, months). Number of months with significant fire activity, which are defined when monthly number of fires is more than 10% of AFD. It is an indicator of seasonality.
3. Yearly standard Deviation of AFD (SDAFD, counts km^{-2} month^{-1}), as an indicator of fire inter-annual persistency.

A dedicated Geographic Information System (GIS) was built with the previously commented fire metrics, as well as with the explicative human factors. The basic grid size was $0.5 \times 0.5^\circ$ resolution, since it provided a good compromise between planetary coverage and data processing (Giglio et al. 2006). The GIS was built only with the cells that include a minimum fire activity, as defined by an average of 0.00004 fire counts km^{-2} month^{-1} during the 6 years of the active fire time series. They totaled 22,381 cells, which cover 31% of the whole emerged land.

14.2.2 Explanatory Variables: Generation of GIS Database

For this analysis a series of explicative variables were derived from standard data sources. The following human variables were included in the analysis (see Chuvieco et al. 2008 for a further explanation): Population density in 2000 (Den2000), Population change between 1990 and 2000 (PopCh), Euclidian distance

(EDPop) to populated places and Cost-friction distances to populated places, considering population density (CostPop). Demographic variables were downloaded from the Center for International Earth Science Information Network (CIESIN) (<http://sedac.ciesin.columbia.edu/gpw/>, last login February 2009). In addition to those variables the Gross Domestic Product per Capita (GDP/Cap) at current US dollars prices was extracted from the International Monetary Fund (<http://www.imf.org/external/data.htm>, last accessed February 2009). Data were only available at country level, but furthermore we computed the GDP by area (GDP/km²), by multiplying GDP/Cap by population density.

In addition to socio-economic variables, land cover categories and Holdridge life zones were included in the analysis to stratify the relationships. Land cover proportions were computed from the Global land cover characterization (v. 2) (<http://edc2.usgs.gov/glcc/>, last accessed February 2009), originally available at 1 km², and proportions of the main land cover types for each half degree cell were computed: Shrubs, Forest, Grass and Crops. The Holdridge Life Zones were downloaded from the UNEP-GRID databases (<http://www.grid.unep.ch/data>, last accessed February 2009) and were generated by the International Institute for Applied Systems Analyses (IIASA) (Leemans 1990). Considering the importance of deforestation processes in fire activity, a forest-crop interface was also calculated. Both the fire metrics and the explanatory variables were referenced to the same 0.5x0.5 grid defined by geographical coordinates.

14.2.3 Statistical Analysis of Input Data

The input database was imported to SPSSTM statistical program (<http://www.spss.com/>) for further processing. Numerical (skewness, kurtosis) and graphical (probability plots) test were computed to test the normal distribution of the input variables, and it was found that most of them greatly depart from normal distribution (Skewness > 3). Therefore, non-parametric statistics were used to measure the relations between fire indices and explicative factors.

Boxplots and graphical exploratory variables (Tukey 1977) were applied to search trends of associations in two (single explicative variable) and three dimensions (two explicative variables).

Kruskal-Wallis one-way analysis of variance was used to test differences in medians of the explicative variables for the quartiles of the different fire indices. The higher the rank differences, the more likely the independent variable is associated to a particular fire index. Tests were run for all explicative variables with AFD, LFP and SDAFD.

The Spearman ρ rank correlation (Healey 1993), was applied to measure the strength of the relationships between AFD, LFP, GDAFD, and all explanatory variables. The coefficient was computed as (SPSS 2006):

$$\rho_s = \frac{T_x T_y - \sum_1^N d_i^2}{2\sqrt{T_x T_y}} \quad (14.1)$$

where d is the rank difference between the two variables for case i ; N the sample size; T_x and T_y are:

$$T_x = \frac{N^3 - N - ST_x}{12} \quad T_y = \frac{N^3 - N - ST_y}{12};$$

and ST_x and ST_y are the sum of ranks for both variables. When T_x or T_y is 0 the statistic is not computed. The significance level was computed from the Student t distribution with $N - 2$ degrees of freedom, following:

$$t = \rho_s \sqrt{\frac{N - 2}{1 - \rho_s^2}} \quad (14.2)$$

Spearman correlation coefficients were computed for the different fire metrics, considering the whole dataset of significant fire activity, as well as for different strata: eco-climatic zones, economic level and land covers. These partial correlations were computed after selecting only those cells that satisfied a specific condition, such as having a dominant boreal climate or being in the lower strata of cropland land cover proportions.

14.3 Results

14.3.1 Spatial Patterns of Fire Indices

Fire density (Fig. 14.1) is more intense in both tropical belts, north and south of the Equator, the former mainly in Africa and, very secondarily, Venezuela-Colombia, and the latter, in Central Brazil, the Congo Basin, North Australia and SE Asia. The impact of sub-Saharan African fires in this regard is enormous, with many areas in the Congo Basin, Gulf of Guinea and Angola-Zambia-Zimbabwe with more than 0.03 fire counts $\text{km}^{-2} \text{month}^{-1}$. Medium to high values can also be observed in Siberia and the Alaska and Northern Canada. A significant belt of fire activity, and with lower values than those mentioned, is also noticeable in SE United States, Central America, a long latitude strip between 45 and 57°N is noticeable in Central Asia (Ukraine, South of Russia and Kazakhstan), and Far East Asia, including the large islands.

Fire seasonal variability is more intense in the Boreal regions, where the dry and snow-free period is short and consequently the fire season too. The opposite can be observed in the subtropical and temperate regions of SE USA, China, Australia, South Africa and Brazil, where fire activity last almost the whole year around. Intermediate values were found in Central Asia and tropical Africa and SE Asia.

Finally, inter-annual variability is larger in those areas where fires affect forested areas or mixed grass-forested, mainly in the Boreal regions of North America and Asia. Additionally, this variable shows high values in the agricultural frontiers

Table 14.1 Kruskal Wallis test for the different fire metrics (all of them have $p < 0.001$)

	Fire density (AFD)	Fire persistency (SDAFD)	Fire seasonality (LFP)
Den2000	272.37	3426.04	1615.07
Popch	431.67	1292.89	917.41
EDPop	793.21	4121.99	2531.84
CostPop	604.48	4101.67	2142.95
GDP/Cap	1532.52	198.27	454.68
GDP/km ²	1552.86	4331.40	1950.65
Shrubs	1354.82	575.62	408.47
Forest	120.95	31.72	934.61
Grass	2747.95	707.17	443.16
Crops	614.90	2595.95	1964.78

of South and Central America, North Australia and scattered regions of Central Africa. Fire persistency, meaning less inter-annual variations is higher in those areas with more agricultural burnings, mainly in SE USA, China and Brazil, most India, Central Asia and the Southern fringe of Australia and Southern Africa, including the eastern side of Madagascar. Table 14.1 shows the results of the Kruskal-Wallis test of the different fire metrics, showing that all explicative variables have significant differences in the four quartiles in which the metrics were divided.

For the fire density (AFD) metric, the variables with higher differences of quartiles were related to economic resources: GDP/Cap and GDP/km², but much lower for those associated to population distribution and population change. The land cover variables did not provided a clear relation with AFD, with the exception of grass and shrub covers.

Ranking differences of the explanative variables with fire inter-annual variability showed highly significant values for distances, population density and GDP/km², and much lower for population change and GDP/cap (all of them $p < 0.0001$). In addition, the ranks are ordered following the order of the quartiles (either positively or negatively), meaning also a consistent trend variation between fire variability and those explanative variables. As far as land cover concerns, the rank differences were lower, but still high values were observed for proportion of croplands.

The rank differences for the length of the fire season (LFP) were significant for most explanative variables, however the ranks followed the order of the LFP quartiles distances, GDP/km², Den2000 and population change, with more significant values for the first three. Land cover proportions offered much lower rank differences and generally they did not follow the LFP quartiles trends, with the exception of the Crop proportions (Table 14.1).

14.3.2 Relations with AFD

Spearman ρ correlations of fire density and explanative variables presented fairly low values. For the whole set of cells (22,381) the higher coefficients were found

for GDP/km² and average Grass cover, the former negative and the latter positive, both highly significant ($p < 0.0001$) (Table 14.2). The correlations within specific strata showed only relevant correlation for the Tropical areas, since both in Boreal and Temperate zones none of the variables reaches ± 0.3 . The areas under forest-cropland or forest-grass interface show higher correlations, especially for those related to economic resources (GDP/Cap and GDP/km², both negative), and distances (EDPop, in this case with a positive sign). For medium to low agriculture proportions, all human explanative variables have low correlations, while in medium to high proportions of agriculture areas, economic and distance values are more important. Same results were observed for Grass proportion, with higher correlations for savanna ecosystems. Forest proportions strata offered generally low correlations, for both high and low proportions, with only medium ρ correlations for GDP/km² and with a negative trend. Finally, the medium to low income areas have higher negative correlation with GDP/km², than medium to high income areas, as well as with distances (in this case, with a positive correlation).

14.3.3 Relations with SDAFD

The Spearman correlation values with human factors were in general higher than for the AFD variable (Table 14.3). The global model provided medium to high correlation ($\rho > \pm 0.4$) for population density and distances (positive) and GDP/km² and cropland (negative). The stratified models showed different trends for each eco-climatic zone. For the Boreal regions, relations of land covers were generally poor, with exception of cropland, and those associated to population density (Den2000, GDP/km²) were higher and negative. In Temperate dry and wet climates the relations were found similar, although less clear than for boreal regions. Interestingly, in these climates, the relation between cropland and fire variability is almost negligible. In Tropical regions, the correlations increased, with medium to high values for most human variables, including population change in Tropical dry, which is not observed in other strata. The land cover groups did not offer much explanation capacity, with the exception of croplands. The highest correlations were observed for EDPop (positive) and Den2000 and GDP/km² (negative). Grasslands and Forest quartile proportions revealed a similar trend as croplands, although generally with higher ρ values. It should be emphasized that similar correlations were observed when only low or only high proportions were considered, meaning the relations are not affected by land cover variation. Again, distances and GDP/km² presented the higher ρ values, while correlations with GDP/Cap and population change are generally very poor. The maximum correlations were observed for the stratification of cells according to GDP/Cap, especially when only rich regions (medium to high quartiles) were considered. The most significant variables are again correlation GDP/km² and Den2000 (with negative signs) and distances (positive sign). In summary, the longer the distances to populated areas, the less population density and the less economic investment by area, the more inter-annual variability of fire.

Table 14.2 Spearman ρ correlations between AFD and human variables. Values for the whole study area (global) and different strata are shown

	Global	Boreal	Tempwet	Tempdry	Tropwet	Tropdry	LowAgri	HighAgri	LowGrass	HighGrass	LowFor	HighFor
Den2000	-0.12	0.01	-0.24	0.03	-0.26	-0.14	0.07	-0.23	-0.11	-0.09	-0.15	-0.08
Popch	0.02	-0.14	-0.20	-0.21	-0.11	-0.07	0.17	-0.07	-0.06	0.08	-0.02	0.07
EDPop	0.20	0.07	0.23	0.16	0.33	0.23	0.03	0.29	0.17	0.22	0.24	0.16
CostPop	0.18	0.02	0.22	0.01	0.33	0.21	-0.02	0.29	0.14	0.17	0.22	0.13
GDP/Cap	-0.23	-0.02	0.00	-0.15	-0.20	-0.19	-0.27	-0.21	-0.07	-0.29	-0.19	-0.27
GDP/km ²	-0.27	0.00	-0.27	-0.09	-0.38	-0.24	-0.07	-0.38	-0.17	-0.33	-0.31	-0.23
Shrubs	-0.22	-0.05	-0.12	-0.05	-0.17	-0.42	-0.25	-0.13	-0.04	-0.36	-0.34	-0.10
Forest	-0.05	0.00	0.01	0.04	-0.07	0.06	-0.15	0.05	0.10	-0.05	-0.06	-0.01
Grass	0.32	0.06	-0.07	-0.03	0.38	0.44	0.37	0.26	0.06	0.31	0.41	0.27
Crops	-0.14	-0.01	0.04	0.17	-0.23	-0.21	-0.04	-0.11	-0.05	-0.20	-0.23	-0.03

Significant $p < 0.001$ correlations are marked in bold. Tempwet and Tempdry are temperate wet and dry climates, respectively; Tropwet and Tropdry are tropical wet and dry climates, respectively; LowAgri and HighAgri are the two lower and higher quartiles of agricultural proportion, respectively; LowGrass and HighGrass are the two lower and higher quartiles of grass proportion, respectively; and LowFor and HighFor are the two lower and higher quartiles of forest proportion, respectively.

Table 14.3 Spearman ρ correlations between fire persistency (SDAFD) and human variables. Values for the whole study area (global) and different strata are shown

	Global	Boreal	Tempwet	Tempdry	Tropwet	Tropdry	LowAgri	HighAgri	LowGrass	HighGrass	LowFor	HighFor
Den2000	-0.41	-0.34	-0.32	-0.17	-0.29	-0.45	-0.32	-0.31	-0.42	-0.37	-0.41	-0.41
Popch	-0.25	-0.21	-0.26	-0.28	-0.15	-0.37	-0.22	-0.21	-0.26	-0.21	-0.26	-0.23
EDPop	0.45	0.38	0.34	0.32	0.35	0.46	0.34	0.37	0.46	0.41	0.47	0.43
CostPop	0.45	0.36	0.31	0.21	0.34	0.49	0.36	0.35	0.46	0.42	0.46	0.44
GDP/Cap	0.04	0.12	-0.04	-0.15	-0.12	0.15	0.09	-0.11	0.05	-0.01	0.04	0.01
GDP/km ²	-0.46	-0.34	-0.39	-0.26	-0.36	-0.46	-0.32	-0.40	-0.47	-0.44	-0.48	-0.43
Shrubs	-0.16	-0.02	-0.03	-0.04	-0.21	-0.20	-0.09	-0.07	-0.15	-0.28	-0.16	-0.15
Forest	0.01	0.14	0.01	-0.05	0.01	-0.08	-0.08	0.03	-0.05	-0.07	-0.17	0.16
Grass	0.18	-0.07	-0.12	0.04	0.28	0.30	0.12	0.18	0.15	0.28	0.30	0.08
Crops	-0.34	-0.34	0.04	-0.03	-0.25	-0.42	-0.29	-0.07	-0.36	-0.34	-0.37	-0.31

Significant $p < 0.001$ correlations are marked in bold

14.3.4 Relations with LFP

The global Spearman correlation for LFP showed very interesting differences according to the strata considered (Table 14.4). The global model showed generally low correlation ($\rho < \pm 0.4$), with higher values distances, although in this case negative, in contrast with AFD and SDAFD variables. Likewise, Den2000, GDP/km² and Crops had positive correlations, again contrary to the other fire metrics.

The stratified models offered remarkable differences for each eco-climatic zone. For the Boreal regions, the correlations of LFP and human variables are very high ($\rho > \pm 0.6$), implying that there is a very significant impact of human factors in these regions. According to the observed relations, fire season tend to last longer where population density and economic resources increase and areas are closer to populated places. In other climatic regions the relations are generally poor ($\rho < \pm 0.3$), especially in Temperate and Tropical dry climates. Tropical wet showed higher values for EDPop, GDP/km² and GDP/Cap, but still close to ± 0.3 .

Stratification by land cover proportions showed interesting differences. For instance, areas with medium to low cropland proportions exhibited medium ρ values for all explanative variables, with the exception of GDP/Cap, while in the medium to high cropland density areas, none was relevant. The same results were observed for Grass proportions, medium to high correlation for low proportions and very low correlation for high proportions, and exactly the opposite for Forest proportions: only those cells with medium to high proportion of forest cover show relevant correlation values. In summary, in areas with low crop and grass and high forest proportions, the impact of the human factors to extend fire season is quite obvious (especially in the last two strata), while in the rest of the areas, human factors are almost irrelevant.

Finally, a clear differentiation was also observed according to the level of economic resources. For poor countries, the relations between fire length and human factors was very weak, while for rich countries, medium to high correlations ($\rho > \pm 0.45$) were computed for most variables. As in other stratifications, the signs were positive for Den2000 and GDP/km² and negative for distances, again implying a positive effect of population presence to extend fire season duration.

14.4 Discussion

This paper provides a first assessment of human patterns and fire parameters at global scale, based on quantitative relations. Although further analyses are required to corroborate the trends identified with our data, an initial set of ideas may be proposed based on the observed results:

Population density and distances to populated places tend to show opposite trends in their relation with fire metrics. Population density is negatively related to fire density and variability and positively to fire length, while distances are positively related to fire density and variability and negatively to fire length. This implies that fires tend to be more frequent, more seasonal and less regular far from populated

Table 14.4 Spearman ρ correlations between fire seasonality (LPI) and human variables. Values for the whole study area (global) and different strata are shown

	Global	Boreal	Tempwet	Tempdry	Tropwet	Tropdry	LowAgri	HighAgri	LowGrass	HighGrass	LowFor	HighFor
Den2000	0.31	0.65	0.20	0.22	0.11	0.06	0.35	0.06	0.49	0.09	0.16	0.48
Popch	0.25	0.12	0.09	0.10	-0.02	0.04	0.38	0.08	0.33	0.10	0.15	0.35
EDPop	-0.38	-0.61	-0.25	-0.28	-0.32	-0.18	-0.40	-0.20	-0.51	-0.26	-0.25	-0.52
CostPop	-0.36	-0.66	-0.23	-0.28	-0.21	-0.12	-0.39	-0.13	-0.52	-0.18	-0.22	-0.52
GDP/Cap	-0.10	-0.24	0.05	0.08	0.31	0.10	-0.17	0.07	-0.18	0.07	-0.01	-0.17
GDP/km ²	0.33	0.65	0.30	0.22	0.33	0.16	0.34	0.15	0.51	0.17	0.20	0.48
Shrubs	0.13	-0.01	0.07	0.08	0.28	0.04	-0.01	0.09	0.13	0.16	0.00	0.17
Forest	-0.19	-0.21	-0.04	0.18	-0.26	0.06	-0.25	-0.04	-0.27	-0.03	0.14	-0.33
Grass	0.16	0.31	0.07	-0.03	0.01	0.00	0.31	0.08	0.13	-0.03	-0.01	0.23
Crops	0.30	0.58	0.04	0.15	0.31	0.15	0.27	-0.03	0.40	0.22	0.15	0.45

Significant $p < 0.001$ correlations are marked in bold

places and in areas where population is less dense. In other words, large populated areas, including the impact of cities, tend to reduce fire occurrence, although the relations are relatively low. Higher correlations are found between human variables and fire persistency, meaning that those areas with more human presence tend to have more similar fire cycles throughout the years than areas with less population presence.

The economic variables used in this study (GDP/km² and GDP/Cap) have very different implications. The latter offers higher relations with fire metrics for most geographical strata. Regularly, the more economic resources available, the more persistent the fires are, and the longer the fire season. This is especially clear in the case of the boreal regions, most probably because of the impact of Siberian fires.

As far as land covers concerns, the presence of grass (savannah areas) tend to imply higher fire density, while the crop areas are negatively correlated. Agricultural areas tend to have more recurrent fires, much less dependent of climatic variations than other land covers, especially in boreal and tropical regions. They are also closely related to the length of the fire season, especially in the boreal regions.

Acknowledgements Data for the global analysis of fire characteristics was downloaded from the following sources: Center for International Earth Science Information Network (CIESIN), International Monetary Fund (IMF), and United Nations Environment Program (UNEP). Efforts of these institutions to generate those databases are greatly acknowledged.

References

- Chuvieco E, Giglio L, Justice CO (2008) Global characterization of fire activity: towards defining fire regimes from earth observation data. *Glob Change Biol* 14:1488–1502
- Cochrane MA, Alencar A, Schulze MD, Souza CM, Nepstad DC, Lefebvre P, Davidson EA (1999) Positive feedbacks in the fire dynamic of closed canopy tropical forests. *Science* 284:1832–1835
- Cortner JH, Gardner PD, Taylor JG (1990) Fire hazards at the urban-wildland interface: what the public expects. *Environ Manage* 14:57–62
- DeFries RS, Houghton RA, Hansen MC, Field CB, Skole D, Townshend J (2002) Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *Proc Natl Acad Sci* 99:14256–14261
- Di Bella CM, Jobbágy EG, Paruelo JM, Pinnock SD (2006) Continental fire density in South America. *Global Ecol Biogeogr* 15:192–199
- FAO (2007) Fire management – global assessment 2006. A thematic study prepared in the framework of the Global Forest Resources Assessment 2005. Rome, FAO Forestry Paper 151
- Giglio L, Csizsar I, Justice CO (2006) Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *J Geophys Res Biogeosci* 111. doi:10.1029/2005JG000142
- Goetz S, Fiske G, Bunn A (2006) Using satellite time-series data sets to analyze fire disturbance and forest recovery across Canada. *Remote Sens Environ* 92:411–423
- Healey JF (1993) *Statistics: a tool for social research*. Wadsworth Pub. Co, Belmont, CA
- Hobbs NT, Schimel DS, Owensby CE, Ojima DS (1991) Fire and grazing in the tallgrass prairie – contingent effects on nitrogen budgets. *Ecology* 72:1374–1382
- Johnson EA, Miyanishi K, Bridge SRJ (2001) Wildfire regime in the boreal forest and the idea of suppression and fuel buildup. *Conserv Biol* 15:1554–1557

- Keeley JE, Fotheringham CJ, Morais M (1999) Reexamining fire suppression impacts on brushland fire regimes. *Science* 284:1829–1832
- Leemans R (1990) Global data sets collected and compiled by the Biosphere Project. IIASA, Laxenburg, Austria
- Leone V, Koutsias N, Martínez J, Vega-García C, Allgöwer B, Lovreglio R (2003) The human factor in fire danger assessment. In: Chuvieco E (ed) *Wildland fire danger estimation and mapping. The role of remote sensing data*. World Scientific Publishing, Singapore, pp 143–196
- Martínez J, Vega-García C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. *J Environ Manage* 90:1241–1252
- Mollicone D, Eva HD, Achard F (2006) Human role in Russian wild fires. *Nature* 440:436–437
- Parisien MA, Peters VS, Wang YH, Little JM, Bosch EM, Stocks BJ (2006) Spatial patterns of forest fires in Canada, 1980–1999. *Int J Wildland Fire* 15:361–374
- Pyne SJ (2001) The fires this time, and next. *Science* 294:1005–1006
- Radeloff VC, Hammer RB, Stewart SI, Fried JS, Holcomb SS, McKeefry JF (2005) The wildland-urban interface in the United States. *Ecol Appl* 15:799–805
- Randerson JT, van der Werf GR, Collatz GJ, Giglio L, Still CJ, Kasibhatla P, Miller JB, White JWC, DeFries RS, Kasischke ES (2005) Fire emissions from C3 and C4 vegetation and their influence on interannual variability of atmospheric CO₂ and D13 CO₂. *Global Biogeochem Cycles* 19. doi:10.1029/2004GB002366
- Reisen F, Brown SK (2006) Implications for community health from exposure to bushfire air toxics. *Environ Chem* 3:235–243
- SPSS (2006) *SPSS Statistics Base 15.0 User's Guide*. Chicago: SPSS Inc
- Tukey JW (1977) *Exploratory data analysis*. Addison-Wesley, Menlo Park, CA
- Vega-García C, Woodard T, Adamowicz, Lee B (1995) A logit model for predicting the daily occurrence of human caused forest fires. *Int J Wildland Fire* 5:101–111
- Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW (2006) Warming and earlier spring increase western US forest wildfire activity. *Science* 313:940–943

Chapter 15

The Use of Remote Sensing Data and Meteorological Information for Food Security Monitoring, Examples in East Africa

Michel Massart, Felix Rembold, Oscar Rojas, and Olivier Leo

Abstract Since 2001, the MARS Unit of the Joint Research Centre of the European Commission has developed a system for crop monitoring and forecasting in food insecure regions. This communication first provides an overall description of the system and then focuses on one monthly bulletin prepared and published by FOOD-SEC action of the MARS Unit in East Africa. The main example is taken from Ethiopia. Basic data, models and information are presented as well as some important parameters for crop monitoring.

15.1 Introduction

Food Security in the poorer countries of the world is a main concern for European Citizens as reflected by the leading role of the European Union in the international donor community with a yearly contribution of around 500 M. Euro. According to the United Nations Food and Agriculture Organization (FAO 2009), 1.02 billion people worldwide are chronically food-insecure.

Like the USAID FEWS Net or the FAO GIEWS initiatives, the FOOD-SEC action of the MARS Unit of the European Commission Joint Research Centre (JRC) has developed since 2001, in cooperation with the European oriented AGRI4CAST action and in the framework of the Global Monitoring for Environment and Security (GMES) initiative, a system for regional crop monitoring and forecasting in various parts of the world. During an initial phase (2001–2004), four pilot areas were covered: Russia and the New Independent States, the Mediterranean Basin, Eastern Africa and South America (MERCOSUR countries plus Bolivia). Since 2005 the capacities of the MARS system has been extended to all the agricultural productive and food insecure areas worldwide.

M. Massart (✉)

EU-DG JRC, Institution for the Protection and Security of the Citizen, MARS Unit TP 266,
Via E. Fermi, 21020 Ispra (VA), Italy
e-mail: michel.massart@jrc.ec.europa.eu

Since 2006, FOOD-SEC action has made specific efforts to strengthen national and regional food security and vulnerability analysis on the Horn of Africa, with additional funding from another European Commission, Directorate General, the DG AIDCO.

The specific objectives of the funding agreement were (1) to improve the quality and coverage of crop monitoring and forecasting activities, (2) to improve the general knowledge and analysis of food security and vulnerability assessment, (3) to support Governments' capacity in implementing the EC funded Food Security Information System projects and (4) to strengthen FOOD-SEC participation in field missions, such as the Crop and Food Supply Assessment Missions of FAO/WFP.

15.2 Region of Interest

The first specific objective is covered by the regular publication of national crop monitoring bulletins on the 6 countries of the Horn of Africa. These bulletins were produced each month during the main crop season for Somalia (since 2001), Ethiopia (2005), Sudan (2005), Eritrea (2006) and Kenya (2007). Similar bulletins are published on South America and on Central Asia. The bulletins are based on meteorological data, satellite remote sensing information and agro-meteorological modelling to analyze crop conditions and to assess crop production.

15.3 Meteorological and Remote Sensing Data

15.3.1 Rainfall

Since 2001, FOOD-SEC is using the meteorological data of the European Centre of Meteorological Weather Forecast (ECMWF) in its bulletins. The data are acquired, pre-processed and delivered every day, every 10 days and monthly to the MARS Unit, under MARSOP, a specific contract with three European companies (ALTERRA, VITO, and METEO Consult). Several parameters, including temperature, precipitation, radiation and potential evapotranspiration are provided at 1 degree grid resolution and, since 2008, at 0.25 degree grid resolution worldwide. Additional indices are also calculated on the initial data, and the latest values are compared with the previous years and with an historical average, on absolute and on relative value. MARS also uses the ECMWF reanalysis data archives "ERA-40" with more than 30 years and ERA-Interim (only 20 years of archive but with a higher spatial resolution), for historical trend studies or to build baseline information for Early Warning Systems.

ECMWF data are meteorological forecasts based on a general circulation model, where all the parameters except precipitation are calibrated in real time with measured data. They are used for their operational status, received worldwide, regularly and in near real time for the bulletin preparation and publication. Nevertheless, they have shown clear limitations, particularly for rainfall information, in some

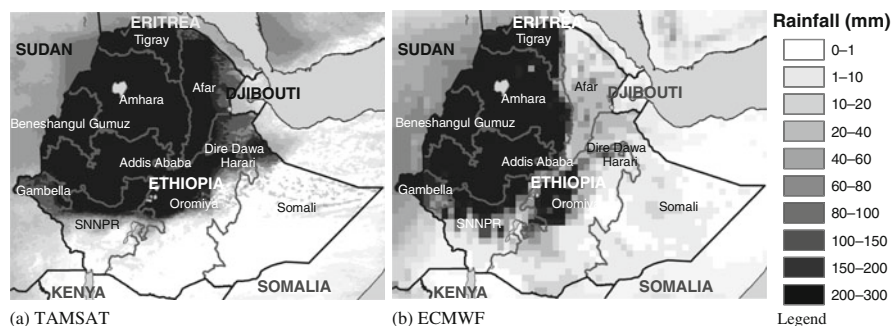


Fig. 15.1 Comparison between TAMSAT (a, left) and ECMWF (b, right) rainfall estimates for the month of August 2009. The two figures show clearly the resolution improvement of TAMSAT approach and the discrepancies between the two model estimates

specific regions. For these reasons, in 2007, FOOD-SEC has decided to invest also in alternative rainfall data. Rainfall estimates based on METEOSAT CCD (Cold Cloud Duration) method are now acquired from the University of Reading (TAMSAT Project). Simultaneously synoptic station data are also collected and archived following a systematic approach at African level. They are complementing the CCD estimations and are integrated in the rainfall estimation process using a geo-statistical approach. The increase of resolution to 5 Km grid represents also a clear advantage of the MSG data compared to the ECMWF information (Fig. 15.1).

15.3.2 Vegetation Condition

FOOD-SEC is using mainly SPOT VEGETATION data in its bulletins. The products received every 10 days worldwide at 1 Km resolution grid, include several standard indices such as the Normalized Difference Vegetation Index (NDVI) and its derivatives: the Vegetation Condition Index (VCI), the Dry Matter Productivity (DMP) and the Vegetation Productivity Indicator (VPI). Monthly composites are also derived. The archive of SPOT VEGETATION data started in April 1998. NDVI difference or anomaly images are computed every 10 days. They show the crop condition difference between the current year and the previous year, as well as between current year and the historical average (of the last 10 years).

Considering the small size of agriculture landscape in some regions, FOOD-SEC has also started to acquire in 2007 MODIS 250 m resolution data over East Africa. The data received from NASA are reprocessed by the Vlaams Instituut voor Technologisch Onderzoek (VITO) to generate ten-daily NDVI composites. MODIS data complement SPOT VEGETATION information for crop monitoring but they are also expected to be used to map the agriculture areas of countries of interest every year.

15.4 The Models

To monitor crop development during the agricultural season two models are currently used by FOOD-SEC experts, the first one is based on meteorological data, the second one uses satellite remote sensing observations.

The FAO Crop specific water balance (CSWB) model is a very simple but physically sound, soil water balance model which is used to assess the impact of weather conditions on crops (Frere and Popov 1979; Gommès 1993; Senay and Verdin 2002, 2003). The water balance of the specific crop is calculated in time increments, usually 10 days. The equation of the water balance is:

$$W_t = W_{t-1} + R - \text{ETA} - (r + i) \quad (15.1)$$

where W_t is the amount of water stored in the soil at the time t , W_{t-1} the amount of water stored in the soil at the end of the previous period ($t-1$), R the cumulated rainfall during the dekad or t -period of time, ETA the actual evapotranspiration in the t -period time, r represents the water losses due to runoff in the t -period time and i represents the water losses due to deep percolation in the t -period time.

Two main outputs of the CSWB model are demonstrated to be positively correlated with the crop yield: the Actual evapotranspiration (ETA) and the Water Requirement Satisfaction Index (WRSI). ETA has the advantage to include the radiation, which is an important climatic variable susceptible to influence the crop yield in the region. The influence of factors other than water stress which can reduce crop yields such as water logging, mechanical damage produced by strong winds, or biological factors, such as locusts, birds, insects or plant diseases are not considered by the CSWB model. The WRSI is an index of the CSWB model to assess the amount of water received by the crop during any time of the season. Normally, the WSI is used for defining qualitative yield classes (i.e. good, average and poor) or in relative figures (percentage of an optimal yield crop). The WRSI expresses the percentage of the maximum crop water requirements which has been met. It is calculated as follows:

$$\text{WRSI} = 100 [1 - (\Sigma|D|/\text{WR})] \quad (15.2)$$

where WRSI = Water Requirement Satisfaction Index expressed in percentage, D = soil water deficit, mm/dekad and WR = maximum plant water requirement, mm/dekad.

The water deficit, D , is set equal to zero whenever W_t from Equation (2) is zero or positive, and D is set equal to W_t whenever W_t is negative (deep percolation and runoff are initially assumed to be equal to zero). The values of D are then totalled and divided by the total seasonal water requirement of the plant to calculate the WRSI. When the WRSI is equal to 100, it indicates no water stress and good crop yields, while a WRSI below 50 corresponds to poor crop yield or crop failures.

The CSWB model (Hoefsloot 2005) included into the FAO AgroMetShell software package has been selected for its simplicity regarding its inputs and for its relevance and reliability in the semi-arid context. A complete description of the model with input and output can be found in Rojas publication (2005). In Addition to the WRSI, specific routines of the CSWB allow to assess the planting date and to follow the development stage of the crops of interest.

The second approach is based on the analysis of the satellite NDVI values during the agricultural season. We used the NDVI index (Rouse et al. 1974), which is the most popular indicator for studying vegetation health and crop production. Research in vegetation monitoring has shown that NDVI is closely related to the LAI (leaf area index) and to the photosynthetic activity of green vegetation. NDVI is an indirect measure of primary productivity through its quasi-linear relation with the fAPAR (Fraction of Absorbed Photosynthetically Active Radiation). The NDVI is extracted and averaged for the crop areas of each administrative unit of interest. For the extraction, an agriculture mask is required. Different techniques can be applied, one of the most advanced being the CNDVI (Crop-specific NDVI) method (Genovese et al. 2001). The method is integrated in FOOD-SEC process when the necessary data are available. The CNDVI extracted every 10 days from the SPOT VEGETATION data gives the profile of the crop development over the season and provides indications on crop performance.

15.5 The Bulletins

Meteorological information, remote sensing data and model output are presented in the monthly bulletins published by FOOD-SEC action. They are usually compared with the previous year and with historical average.

The first page of the bulletin (Fig. 15.2) presents the general agriculture monthly conditions of the country, displaying the Vegetation Condition Index of the current month with a short text commenting on the season and an overall assessment of the evolution of the season with a green to red sign.

The second page shows the rainfall conditions of the month, rainfall of each 10 days of the month and the monthly rainfall compared to last year, to the average of the last 10 years or to an historical average of the last 30 years. Cumulative rainfall graphs from the beginning of the season are also sometimes displayed for specific agro-meteorological regions or watersheds.

The following pages of the bulletin display crop specific information. The most important staple crops of the country are selected. The CSWB outputs maps (planting date, crop cycle progress or qualitative estimated yield), the NDVI monthly difference maps with last year or with historical average and graphs of specific NDVI profiles are presented. The profiles are displayed for the main productive administrative units which have been selected to represent at least 50% of the national crop production. The graphs not only show the NDVI profile of the current season compared with last year and with an average but also the rainfall profile

FOOD-SEC
Crop monitoring in Ethiopia
June 2008

Date of issue: 12 July 2008 Vol. 01 - 2008

Fair start of the season in the Central Highlands of Ethiopia

During the month of June, heavy rains have been observed in the western part of Ethiopia. These rainfall conditions are nevertheless below normal and below last year condition for most of the main cereals production areas. The Vegetation Condition Index map on the right of this page indicates that most of the country is in one of the good year of the last nine years. A low VCI area appears nevertheless in Eastern Amhara, and can be of real concern during the planting period in the production zones of North and South Welo. Following an unfavourable BELG season, the VCI map reports again difficult conditions in the pastoral areas of south-east Oromiya, Somali and Afar regions where the rainfall has been below average and below last year for the last months.

Figure 1. Vegetation Index of June 2008 in relation to the historical maximum and minimum values of the period 1999 to 2007

Produced by EC - JOINT RESEARCH CENTER
IPSC (www.ipsc.eurc.org) FOOD-SEC (www.foodsec.org) and FAO
Contact: Head of WARS Unit, Jacques Dubois - FOOD-SEC Coordinator - Chantal Laroche
Benoit Anquetin - Michel Rueland - Nicolas Verdier - Jean-Louis Verneer

JRC
EUROPEAN COMMISSION

ECMWF MONTHLY WEATHER - June 2008 - FOOD-SEC - Joint Research Center - EUROPEAN COMMISSION

Rainfall analysis

Figure 2. Dekadal rainfall during June 2008. Data derived from ECMWF model.

Figure 2 shows the dekadal rainfall in June according to the ECMWF model. Heavy rains are observed in the Western part of Oromiya and in the Beneshangul regions, mainly for the first and second decades of the month. The rains remain particularly low for the season in the Amhara and Tigray regions.

Figure 3. Monthly June rainfall anomaly difference in mm : with long term average (1994-2003) on the left and with last year (2007) on the right

The difference with last year rainfall and with the long term (LT) average rainfall for June 2008 presents different patterns instead to the regions. Rainfall conditions are clearly below last year and below average for Tigray, Amhara, Beneshangul and SNNPR regions. In Oromiya, the comparison with the LT average shows also below average conditions but nevertheless, in the Central Highlands, rainfall conditions appear to be still above last year.

2

ECMWF MONTHLY WEATHER - June 2008 - FOOD-SEC - Joint Research Center - EUROPEAN COMMISSION

Remote Sensing Analysis - Agriculture - Maize

The monthly NDVI difference with historical average (1999-2007) for June shows normal or slightly above average values located mainly in Central Ethiopia. Below average spots affect the southern part of the country and specifically the maize production areas of South SNNPR and South Oromiya. Looking at the vegetation index profiles of the cultivated areas (Fig. 6) for Maize, above last year and above average NDVI profiles are observed in all the zones of interest, except in Liben where the NDVI profile is since May below normal. Maize is the main cereals in Liben zone where the BELG season was reported as clearly below last year and a complete failure in some areas. During June, the rainfall has been below average and below last year in Central Amhara and in SNNPR region. This may strongly affect maize planted end of March in SNNPR region.

Figure 4. NDVI Monthly difference with the long term average (1999-2007) and location of the zones of interest of Fig. 6.

Figure 5. Estimated Maize planting date (CWS model) for the main maize production areas

3

ECMWF MONTHLY WEATHER - June 2008 - FOOD-SEC - Joint Research Center - EUROPEAN COMMISSION

Figure 6. Time series of crop specific NDVI and rainfall patterns in Ethiopia. The graphs compare the current season with the historical average and the previous year.

4

Fig. 15.2 (a-d) First pages of the Ethiopian bulletin with the country crop conditions, the rainfall information and the specific crop condition pages

of the current season compared to the average rainfall. The two sets of information are complementary. Maps and graphs are commented to evaluate the potential crop yield at the end of the season.

The crop condition assessments are only qualitative during the first part of the agricultural season, comparing the profiles or the maps of the current season with previous year or with historical average and identifying hot-spots of concern (Fig. 15.3).

When possible, quantitative forecasts have been developed using various methods: in Somalia, Sudan, the Mediterranean Basin and in Central Asia, a three steps approach has been introduced in 2005 (Fig. 15.4). The first step gives a broad estimate of the production or the yield by looking at the historical trend and by identifying possible periodicity. The two further refinement steps are based on crop profile similarity analysis with previous years and on regression computation between national statistics and several crop status indicators. The final evaluation is then made by an expert, taking into consideration the results of the three steps.

For Eritrea, yield estimates of cereals have been developed in 2008, based on a simple linear regression model between NDVI (integrated over the season) and the official agricultural statistics. Specific regression was calculated for each of the four main agricultural regions using the data of the last 10 years, with an average coefficient of determination of the regression of 0.70 (R^2) (Fig. 15.5).

For Kenya, a multivariate model has been developed in 2006, involving remote sensing NDVI in combination with Actual Evapo-Transpiration (ETA) calculated

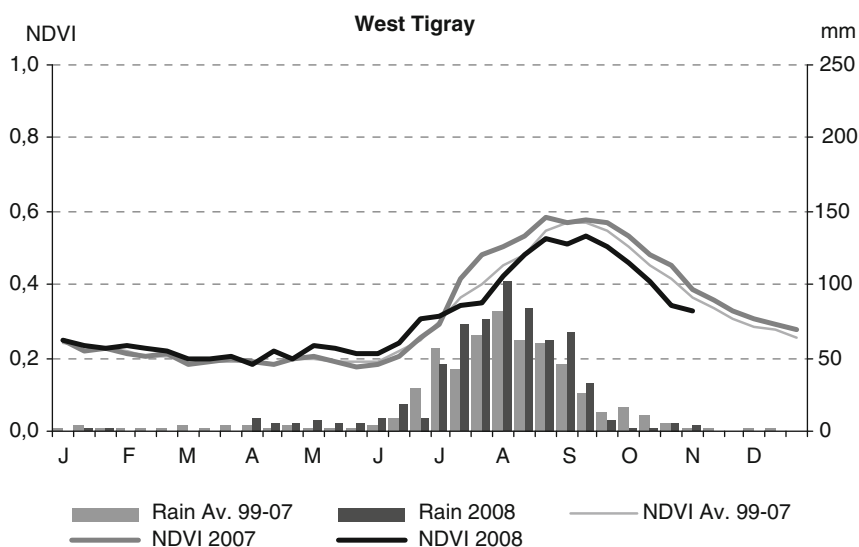
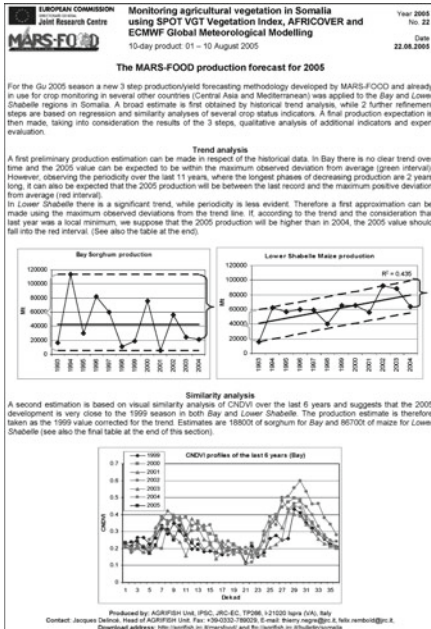
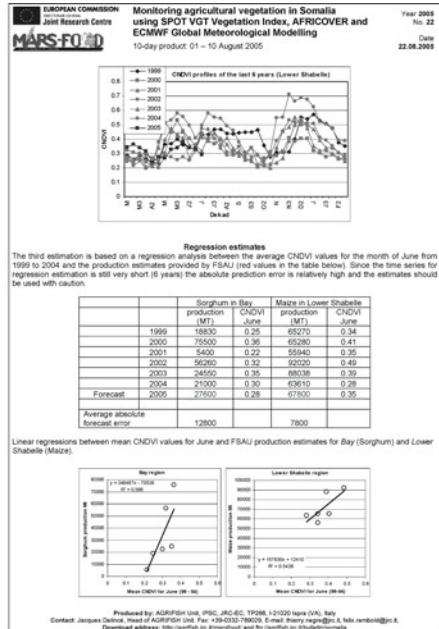


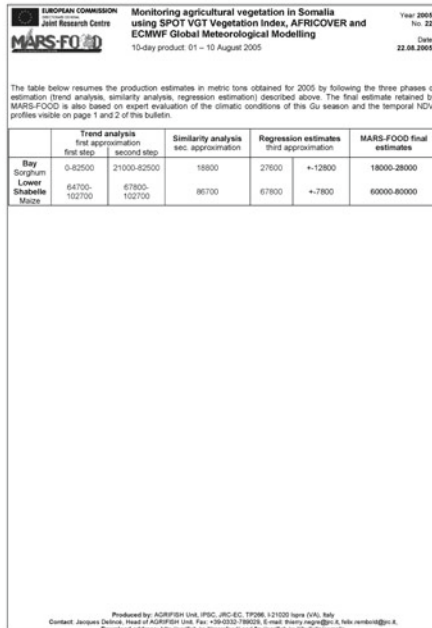
Fig. 15.3 Crop profiles analysis presented in the bulletins. An example for West Tigray, Ethiopia. Forecast for the crops of the 2008 Ethiopian MEHER season: expected yield is below average and previous season



(a)

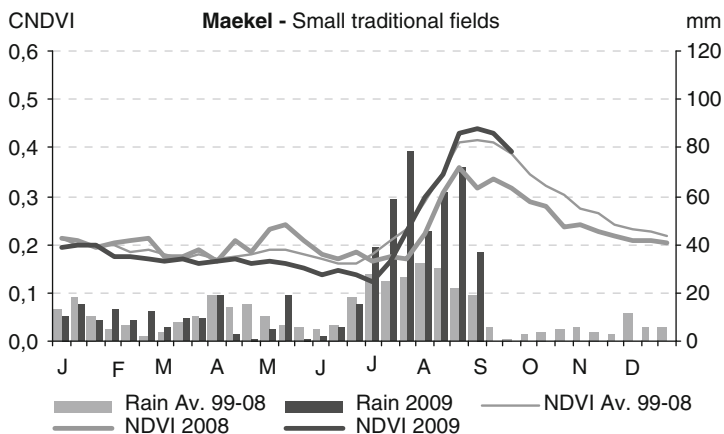


(b)

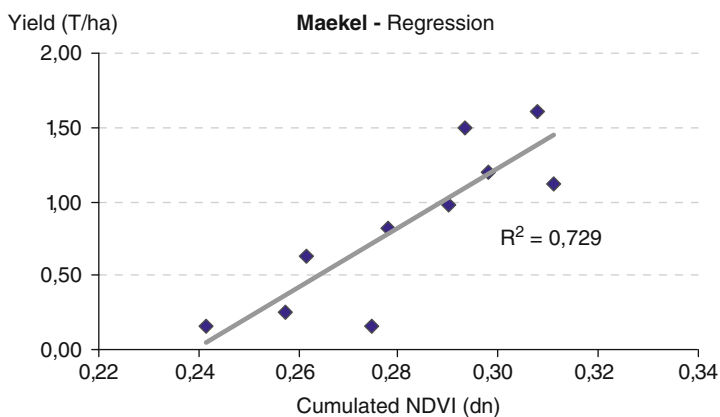


(c)

Fig. 15.4 (a-c) Three steps quantitative approach used in Somalia in 2005: trend analysis, similarity and regression approach, final expert assessment



(a)



(b)

Fig. 15.5 (a and b) Crop profiles and cereals regression model for Maekel region of Eritrea in 2009

by the CSWB model in the yield estimation (Rojas 2007). A coefficient of determination (R^2) of 0.80 has been reached with this model. Using the regression model FOOD SEC is estimating the maize production during the “Long Rains” crop season in Kenya since 2006 (Fig. 15.6). Accuracy is assessed retroactively comparing FOOD SEC forecasts with Governmental Statistics.

The quantitative yield forecast is a challenging research activity which can only become operational under certain conditions such as reliable and detailed agricultural statistics, precise crop phenology information and good quality estimators. In any case, the approach can not directly include factors such as weeds, pests, conflicts, etc . . . In this context, field data and close collaboration with local institutions are crucial for monitoring and especially forecasting activities.

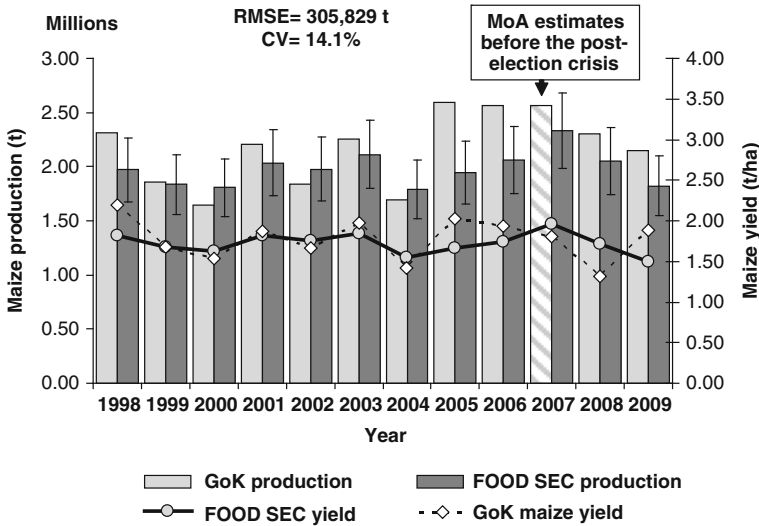


Fig. 15.6 Results of the forecasting model for Long Rain’s maize production (1998–2009) in Kenya. Observed productions are the official statistics from the Government of Kenya (GoK). Note: 2007 figure was done before the post-election crisis; 2008 and 2009 figures still provisional. Note that the yield is the main factor for the maize production estimate in Kenya. This is due to the fact that the maize area has a much smaller inter-annual variability ($CV = 7.4\%$) compared to the 27% inter-annual variation of yield

15.6 Some Important Parameters

15.6.1 Agriculture – Crop Mask

To monitor crop development, satellite information, such as NDVI, is specifically extracted on crop areas. In East Africa, the AFRICOVER project produced land cover maps for several countries in early 2000. The land cover maps developed using the LCCS classification system (Di Gregorio and Jansen 2000) request a class interpretation to derive a crop mask. The resolution of the final maps produced from visual interpretation of LANDSAT data is higher than the SPOT VEGETATION data which allows the calculation of a weighted crop specific NDVI or CNDVI. This approach is applied in Sudan, Eritrea, Kenya and Somalia.

Ethiopia has not been covered by the AFRICOVER project (AFRICOVER 1995). A national land cover map developed by the World Bank and the Ministry of Agriculture and Rural Development project, the “Woody Biomass Project”, is available since 2000. The map is also based on LANDSAT interpretation but the resolution is not fully consistent at national level. To refine the information and to reach a “more” crop specific map, the “cultivation” class of the woody biomass map has been crossed with elevation data which conditions the distribution of the crops in Ethiopia (Reynolds 2007).

In the absence of any high resolution agriculture map, the Global Land Cover 2000 product has been used as agriculture mask. The map which is available worldwide has been produced in 2000 based on the classification of SPOT VEGETATION data at 1 Km resolution. In the same context, other mid resolution products based on MODIS or MERIS data are currently available, developed under various initiatives like the GMFS program or the ESA-GLOBCOVER project, they show a clear improvement to GLC 2000 in terms of resolution. The precise accuracy of the agriculture class still has to be assessed. However, the advantage of using an agriculture mask for extracting the NDVI is highlighted by a comparison between NDVI extractions with or without mask (Fig. 15.7). The mask improves the quality of the information with the focus on agriculture. Nevertheless, the difference between the use of a high resolution (LANDSAT based) and a mid-resolution mask (SPOT VEGETATION or MODIS) is not always straightforward, at least for a qualitative evaluation.

15.6.2 Crop Phenology

The importance of rainfall in the CSWB model is clear, as it is the main input, and is discussed in the paper of Rojas (2007). Another element of the CSWB model is the crop phenology and in particular the planting/sowing date. Planting date windows (average planting period) are available from FAO databases like the IGAD “Crop Production System Zone” database (Van Velthuisen et al. 1995), or often from national sources. But the relevant input for running properly the CSWB model is the planting date of any given year. This information can be observed on the ground. The National Meteorological Agency (NMA) in Ethiopia is reporting crop phenology during the cropping season for around 30 meteorological stations. This information is valuable to calibrate the model but is unfortunately often not complete or available for the whole season. Another approach is to derive the planting/sowing date from models based on rainfall patterns or NDVI vegetation profiles. FOOD-SEC is currently using a model based on the accumulation of significant rainfall during three (one with 10 mm rain plus 2 with a sum of 20 mm) dekads at the beginning of the rainy season. The parameters of the model derived from rules defined in the Sahel region (AGRHYMET Centre) have been adjusted for East Africa context. VGT4AFRICA program and JRC provide also a “vegetation start of the season” product based on NDVI profiles analysis. The product is not specifically a “crop planting or sowing” date product but it could be used as a proxy. It thus needs to be calibrated and validated for crops.

15.6.3 National Agriculture Statistics

The production of quantitative yield estimates needs the calibration of the model with historical crop yield statistics. Specific attention must be paid on the spatial

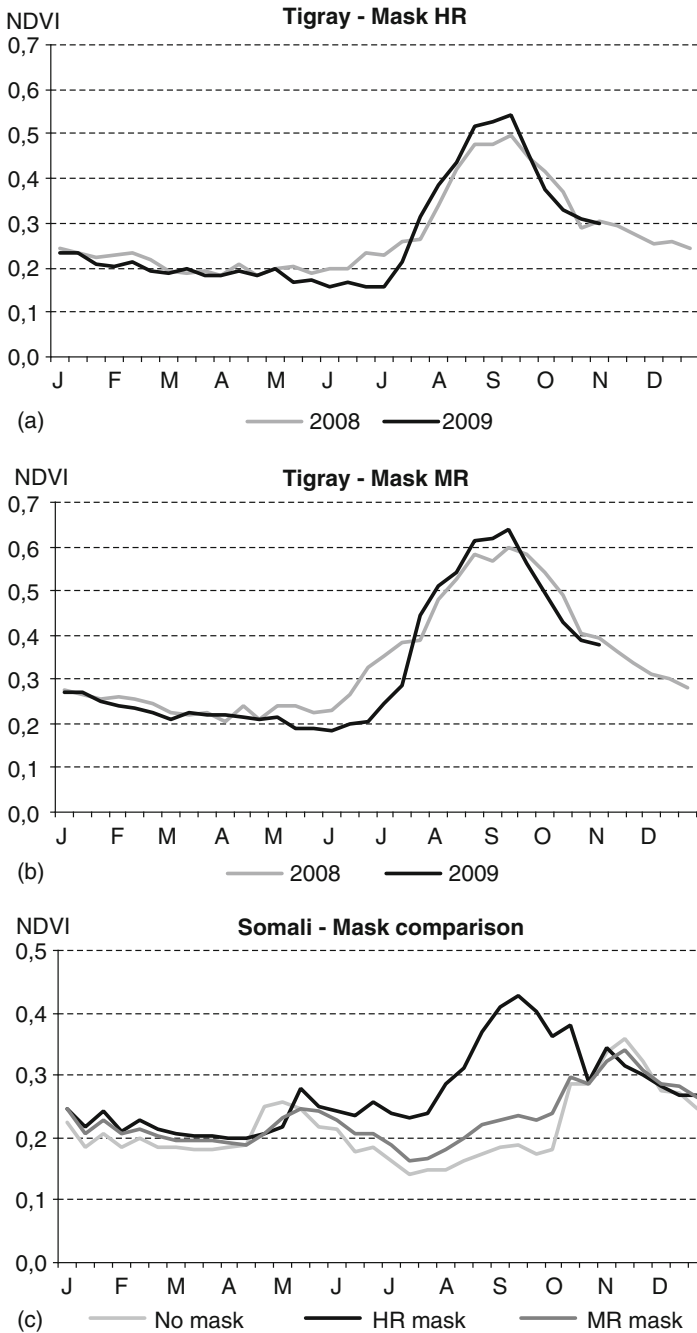


Fig. 15.7 (a–c) Comparison between NDVI profiles calculated with different agricultural masks in Ethiopia. In Tigray region, the use of a high resolution mask versus a mid resolution one does not show a clear improvement for inter-annual comparison. In Somali region, the interest of the HR mask is highlighted. The HR mask allows a good identification of the Gu crop season (August–October) compared to a MR mask and to the absence of mask

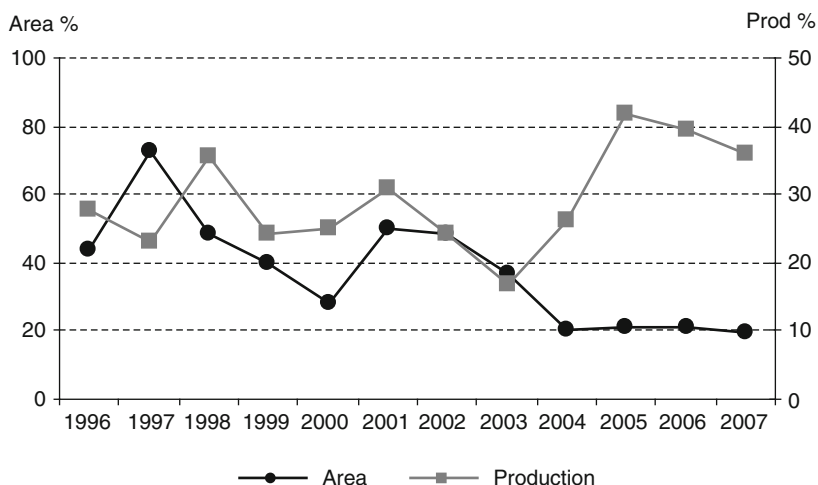


Fig. 15.8 Differences of area and production of cereals and pulses between the agricultural statistics published by the Ministry of Agriculture-CFSAM and the CSA in Ethiopia (source SFSIS, FAO). The differences are expressed in relative values (percentage)

level at which these data are available to be able to consider homogeneous areas, as well as on the length of the time series of the statistics. FAO is providing agriculture statistics but usually at national level or for large administrative units. To get better spatial information, in-country collection is needed. The quality of the model estimation relies of course on the quality of these statistics which is not always easy to assess. In Ethiopia, two sources of agriculture statistics are available, from the Central Statistical Authority (CSA) and from the Ministry of Agriculture (Fig. 15.8). They are produced using two different field survey methods. Since 2006, an EU funded project implemented by FAO is trying to reconcile the two approaches. For their time consistency and easy availability, FOOD-SEC is using today the CSA data in Ethiopia.

15.7 Field Assessment

At the end of the agriculture season, FOOD-SEC team is involved in national crop assessment surveys where the FOOD-SEC products, bulletins and estimations can be compared to ground observations.

FOOD-SEC team participates as EU Observers, to Crop and Food Supply Assessment Missions (CFSAM) carried out by FAO and WFP at the request of the Governments at the end of the main agriculture season.

In Ethiopia, since 2006, FOOD-SEC is also joining USDA experts for a Crop Assessment Tour before the harvest to evaluate crop development. Information is collected by interviewing farmers and observing fields in the main agriculture

regions of the country. During the Tour, field photographs are taken regularly and geo-referenced with a Global Positioning System to be further related to satellite images or agro-meteorological outputs.

15.8 Data Dissemination

FOOD-SEC bulletins are produced 7–15 days after the end of each month and disseminated by email using specific distribution lists per country. Each bulletin reaches directly around 100 persons; European Commission officers, EU Delegations, National institutions of the country concerned, research centres and UN institutions (FAO and WFP). The bulletins are also available on line on the INTERNET site of the MARS Unit: <http://mars.jrc.ec.europa.eu/mars/About-us/FOODSEC>.

The FOOD-SEC data maps are available on line, on the specific INTERNET site of the MARSOP project (Fig. 15.9). On the site, NDVI temporal profiles for specific regions of interest can be calculated. The maps are available at: <http://www.marsop.info/>.

The FOOD-SEC meteorological data used in the bulletin can be extracted following different format for specific regions of interest on: <http://cidportal.jrc.ec.europa.eu/home/idp/thematic-portals/foodsec-imageserver/>.

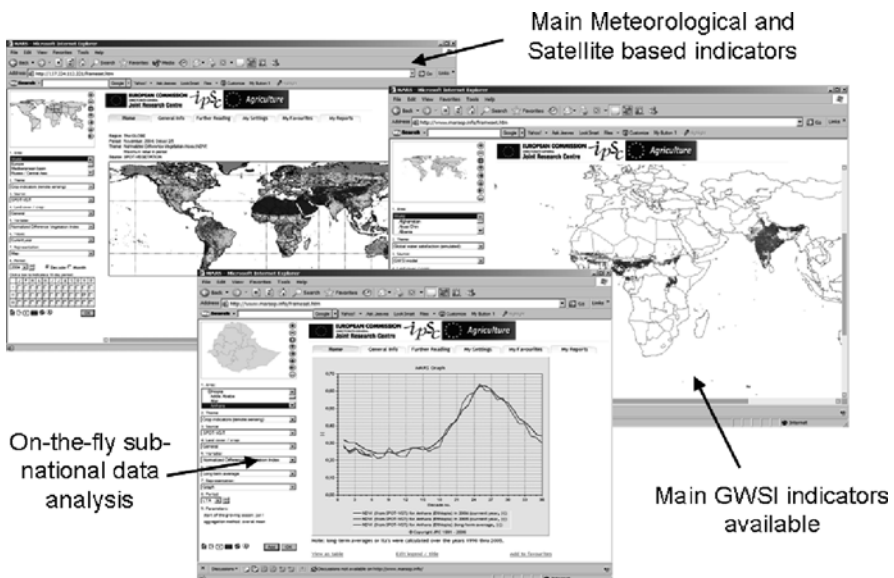


Fig. 15.9 Examples of MARSOP site products and data dissemination procedure

15.9 Research and Development

FOOD-SEC conducts various research and development activities to improve the crop monitoring systems and better understand the hazards that can affect food security. For instance:

- Analyze the trends observed on AVHRR NDVI long time series to assess the potential impact that climate change could have on African agriculture (Vrieling et al. 2009).
- Study on assessing drought probability and severity of impact on agriculture in Africa based on remote sensing data (Rojas et al. 2009a,b).
- Rainfall estimates derived from satellite imagery and global circulation models are frequently used for vegetation monitoring in many areas of Africa, due to the shortage of observed rainfall data and the sparse network of meteorological stations. At the same time, this scarce density of rain gauge stations makes the calibration and validation of the modelled data nearly impossible. In this study, a methodology was proposed for a rapid quality assessment of rainfall estimates, which is based on the well known relationship between rainfall and NDVI (Rojas et al. 2009a,b).

References

- AFRICOVER (1995) AFRICOVER project of United-Nations Food and Agriculture Organization. <http://www.africover.org/>. Accessed on May 2010
- Di Gregorio A, Jansen L (2000) Land cover classification system (LCCS). Classification concepts and user manual. FAO, Rome
- FAO (2009) The state of food insecurity in the world. Economic crises, impacts and lessons learned. FAO, Rome
- Frere M, Popov G (1979) Agrometeorological crop monitoring and forecasting. FAO Plant Production and Protection Paper No. 17. FAO, Rome
- Genovese G, Vignolles C, Negre T et al (2001) A methodology for a combined use of Normalized Difference Vegetation Index and CORINE land cover data for crop yield monitoring and forecasting. A case study on Spain. *Agronomie* 21:91–111
- Gommes R (1993) FAOINDEX Version 2.1. Agrometeorology Group. Environment and Natural Resources Service, SDRN. FAO, Rome
- Hoefsloot P (2005) Agrometshell Version 1.5. Agrometeorology Group, Environment and Natural Resources Service, SDRN. FAO, Rome
- Reynolds C (2007) Crop Tour report for Niger, Burkina Faso and Ethiopia, FAS-USDA. http://www.pecad.fas.usda.gov/highlights/2007/12/Ethiopia_BF-Niger/. Accessed on May 2010
- Rojas O, Rembold F, Royer A et al (2005) Real-time agrometeorological crop yield monitoring in Eastern Africa. *Agronomie* 25:63–77
- Rojas O (2007) Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. *Int J Remote Sens* 28(17):3775–3793
- Rojas O, Vrieling A, Rembold F (2009a) Assessing drought probability for agricultural areas in Africa with remote sensing. Submitted to *Remote Sens Environ*
- Rojas O, Rembold F, Delincé J et al (2009b) Using NDVI as auxiliary data for rapid quality assessment of rainfall estimates in Africa. Submitted to *Int J Remote Sens*

- Rouse JW, Haas RW, Schell JA, Deering DH, Harlan JC (1974) Monitoring the vernal advancement and retrogradation (Greenwave affect) of natural vegetation. Greenbelt, MD. USA: NASA/GSFC
- Senay G, Verdin J (2002) Evaluating the performance of a crop water balance model in estimating regional crop production. In: Proceedings of the Pecora 15 Symposium, Denver, CO
- Senay G, Verdin J (2003) Characterization of yield reduction in Ethiopia using a GIS-based crop water balance model. *Can J Remote Sens* 29(6):687–692
- Van Velthuisen H, Verelst L, Santacroce P (1995) Crop production system zones of the IGADD sub-region. Agrometeorology Working Paper Series No. 10, FAO Rome
- Vrieling A, de Beurs KM, Brown ME (2009) Phenological characterization and variability of African farming systems with NDVI time series. Submitted to *Climatic Change*

Chapter 16

Application of an Early Warning System for Floods

Adriana Albanese, Piero Boccardo, Fabio Giorgi, Nishshanka Prasanna Premachandra, Olivier Terzo, and Rossella Vigna

Abstract Nowadays real time flood forecasting is becoming an important issue as climate changes make quite inefficient those traditional flood forecasting methods based only on historical data. Furthermore, real time flood forecasting systems are often limited to developed countries and difficult to find in developing countries where more than three-quarters of the world's population are living. In this context, the development of an early warning system for floods able to provide global coverage is very important. The aim of this paper is to show the improvements introduced in the Early Warning System for Floods, developed by ITHACA organization, i.e. mainly the calculation of those morphometric parameters necessary to understand the basin response to heavy rainfalls and the source Real Time data adjustment necessary to run the detection of critical rainfalls in near real time. This completely automated System runs in river basin scale having a global coverage by using 3B42 and 3B42RT satellite rainfall data products of Tropical Rainfall Measuring Mission (TRMM). The 3-hourly 3B42 data from 1998 to 2007 are used to detect all historical flood events in the past 10 years using a hydrological method based on Depth Duration Frequency curves. The 3-hourly real time 3B42RT data with some statistical adjustments are used to detect critical rainfall events and to make alerts in near real time.

Abbreviations

CRED	Centre for Research on the Epidemiology of Disasters
DBMS	Database Management System
DDF	Depth Duration Frequency
EM-DAT	Emergency Events Database
EROS	Centre for Earth Resources Observation and Sciences
ETL	Extraction, Transformation and Loading
FEWS NET	Famine Early Warning Systems Network

A. Albanese (✉)
DITAG, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy
e-mail: adriana.albanese@polito.it

GLIDE	GLobal IDentifier number
IDL	Interactive Data Language
ITHACA	Information Technology for Humanitarian Assistance, Cooperation and Action
OCHA	UN Office for the Coordination of Humanitarian Affairs
RT	Return Time
SDI	Spatial Data Infrastructure
SiTI	Istituto Superiore sui Sistemi Territoriali per l'Innovazione
TRMM	Tropical Rainfall Measuring Mission
TMPA	Multisatellite Precipitation Analysis
USAID	United States Agency International Development
WFP	World Food Programme

16.1 Introduction

Climate projections using multi-model ensembles show increases in globally averaged mean water vapour, evaporation and precipitation over the 21st century. The models suggest that precipitation generally increases in the areas of regional tropical precipitation maxima (such as the monsoon regimes, and the tropical Pacific in particular) and at high latitudes, with general decreases in the sub-tropics. Widespread increases in heavy precipitation events (e.g., above the 95th percentile) have been observed, even in places where total amounts have decreased. The observed increase in precipitation intensity indicates that climate change might already have had an impact on the intensity and frequency of floods. Heavy precipitation events are projected to become more frequent over most regions throughout the 21st century (Bates et al. 2008).

Because of the global climate change, the floods induced by storm events are becoming more frequent than in the past; furthermore, floods are the major cause of personal injuries and proper damages and the problems related to flooding have greatly increased over decades because of population growth and the subsequent development of extensive infrastructures in close proximity to rivers (Al-Sabhan et al. 2003).

Since flooding is one of the most disastrous hazards in many regions of the world (Schanze 2009) and it is increasing in the last years, the flood risk management has become very relevant. However, real time flood forecasting systems are often limited to developed countries and difficult to find in developing countries where more than three-quarters of the world's population are living. In this context, the monitoring of rainfall and the development of an early warning system for floods able to provide global coverage is very important.

This paper reports the description and the application of an Early Warning System for Floods based on monitoring of rainfall on worldwide scale. A description of the structure and methodology, parameters and data requirements of the system and the results obtained are also presented. This project is conducted by

Information Technology for Humanitarian Assistance, Cooperation and Action (ITHACA, www.ithacaweb.org) which is a non-profit organization founded by Politecnico di Torino and SiTI (Istituto Superiore sui Sistemi Territoriali per l'Innovazione, www.siti.polito.it). ITHACA conducts research activities in the field of Geomatics and particularly it devotes its activities to monitor, analyze and forecast natural disasters. This system was developed under a special request made by World Food Programme (WFP, www.wfp.org) in order to increase efficacy in approaching emergency preparedness related to flood events. The aim of this Early Warning system is to give an alert in advance about the occurrence of floods around the world monitoring the heavy rainfalls in near real time; this system can be used by WFP or other humanitarian assistance organizations to evaluate the events and to understand the potentially floodable areas where their assistance is needed.

16.2 The Structure of the Project

This Early Warning System is based on precipitation analysis and it uses rainfall data from satellite at worldwide extent.

The first step in this project was the individuation of the different extreme meteorological events occurred in the past years using a hydrological analysis and the calculation of cumulated rainfall values during these events. To produce this model, it has been necessary to define a specific morphometric parameter called lag time, evaluated for each river basin (see Section 16.3.2).

The same analysis is performed using near real time data to monitor current rainfall conditions every 3 h. This can be considered the proper Early Warning System for flood events (see Section 16.4).

16.2.1 Data

The rainfall dataset used in this project belong to Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA). The TMPA provides a calibration-based sequential scheme for combining precipitations estimates from multiple satellites, as well as gauge analyses where feasible (Huffman et al. 1997, 2007); in particular, in this project only two products have been used: the 3B42 and the 3B42RT (see Section 16.2.2). These gridded estimates are on a 3-h temporal resolution and 0.25° by 0.25° spatial resolution in a global belt, extending from 50° South to 50° North latitude.

The analysis of rainfall data are led on the basis of river catchment due to the strict correlation between the catchment characteristics and flood effects. A hydrological catchment can be considered as the geographical surface area and geologic subsurface structure which delivers water to each trunk river. This three-dimensional land system is bounded by a watershed (Smithson et al. 2002). The GIS watershed layer of HYDRO1k, developed at the U.S. Geological Survey's Centre for

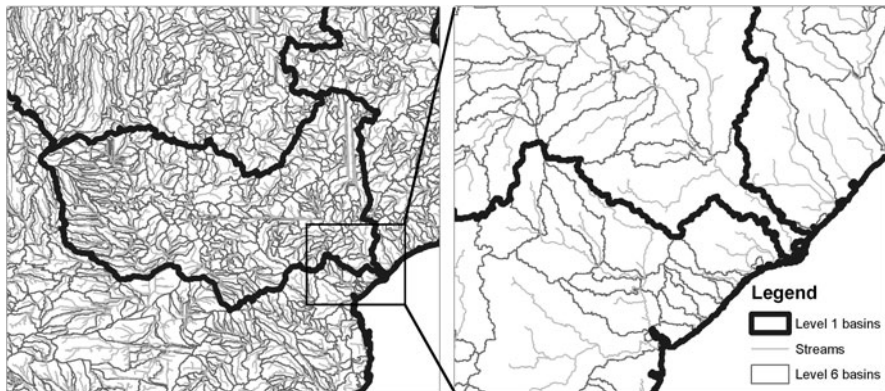


Fig. 16.1 Stream line and drainage basin layers from HYDRO1k. The watershed layer is constituted by a territorial subdivision from level-1 to level-6, after the HYDRO1k criterion

Earth Resources Observation and Sciences (EROS) has been used in this project (<http://eros.usgs.gov/products/elevation/gtopo30/hydro/index.html>). HYDRO1k is a geographic database developed to provide global coverage of topographically derived datasets, including streams, drainage basins and ancillary layers derived from a digital elevation model of the world (GTOPO30). This layer of watersheds is constituted by a territorial subdivision at different levels, increasing details from level-1 to level-6 (Fig. 16.1); watershed at maximum detail (level-6 basin) is here taken as reference for the system.

16.2.2 Pre-elaboration

A pre-elaboration of rainfall data is required to work at the granularity of level-6 basin. The rainfall data are originally associated to latitude and longitude values and then reconducted to a new territorial entity, called `geodb_id`. The `geodb_id` is the result of the intersection between the TRMM grid and the HYDRO1k watershed layer (Fig. 16.2), in other words it is the portion of level-6 basin area covered by the TRMM grid cell ($0.25^\circ \times 0.25^\circ$ latitude and longitude). Once this pre-elaboration is conducted the area weighted average rainfall is calculated for each level-6 basin and it is the value used by the hydrological model to calculate critical rainfalls.

It has been decided to use Oracle 10 g[®] to load and elaborate the whole dataset of rainfall that was structured in a data warehouse architecture (Fig. 16.3). This allows the management of this huge dataset by performing complex data analyses and queries.

An other pre-elaboration of data is required to transform data into a suitable format for the database. The primitive raster format of the gridded data TRMM 3B42 and TRMM 3B42RT was transformed into the ASCII format, to be loaded into Oracle. Both pre-elaboration procedures are performed by an ITHACA routine elaborated in a suitable environment for the raster elaboration, the Interactive Data Language (IDL).

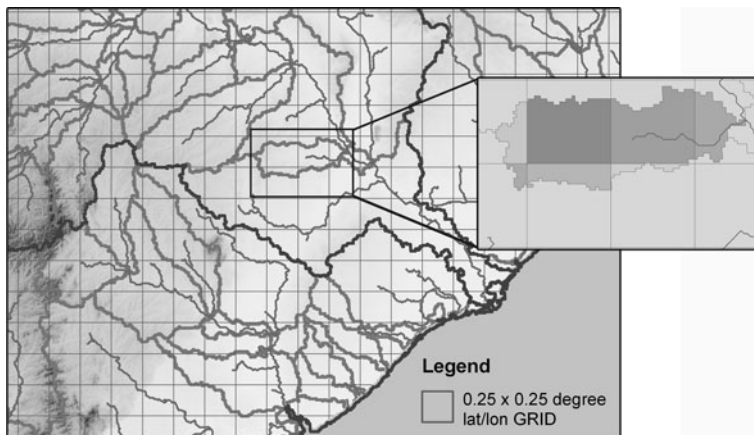


Fig. 16.2 Intersection between the HYDRO1k watershed layer and the TRMM grid; each basin contains different new territorial entity that are the geodb_ids (in gray scale in the map)

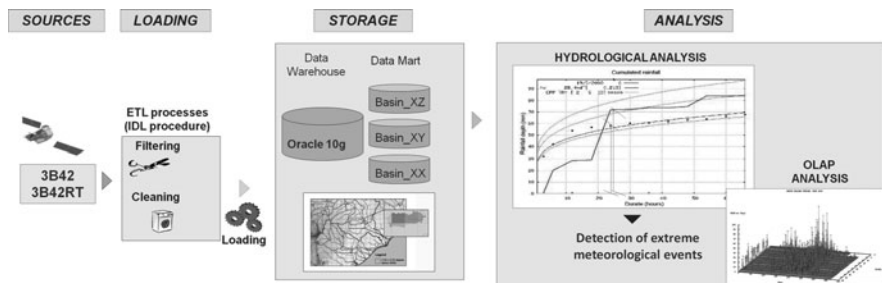


Fig. 16.3 Early warning system architecture

The data model has been implemented with the star schema concept (Golfarelli and Rizzi 2002), where rainfall data were stored in different tables considering two levels of dimensional aggregation: the spatial and the temporal reference. Spatial reference is defined with the georeferenced information from level-1 to level-6 basins. Time reference is defined with the year, month, day and 3'h fields (Albanese et al. 2008).

Once the data are loaded they become ready for the different analyses. In particular, the first analysis allows the detection of all historical flood events occurred in the past (see Section 16.3); subsequently, the whole system runs in near real time in order to find current critical rainfalls (see Section 16.4).

16.3 Analysis of Historical Data

The individuation of extreme meteorological events occurred in the past was executed elaborating the whole dataset of 3B42 product (<http://trmm.gsfc>).

nasa.gov/3b42.html), belonging to the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA); it covers the period from 1998 to the delayed present.

The 3B42 product merge high quality (HQ)/infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates. This dataset is a post-real time research-quality product and is available about 10–15 days after the end of each month (Huffman et al. 2007). The data are produced in four stages: (1) the microwave precipitation estimates are calibrated and combined, (2) infrared precipitation estimates are created using the calibrated microwave precipitation, (3) the microwave and IR estimates are combined, and (4) rain gauge data are incorporated. It is highly advantageous to include rain gauge in combinations datasets (Huffman et al. 1997).

The hydrological analysis (see Section 16.3.1) is developed for the elaboration of 3B42 data, in order to individuate critical rainfalls occurred in the last 10 years (from 1998 to 2007).

16.3.1 Hydrological Analysis

Rainfall data have been elaborated using the hydrological method of Depth Duration Frequency (DDF) curves and the morphometric parameter called lag time (see Section 16.3.2) has been considered to create a time scale where critical durations had to be detected.

Depth Duration Frequency curves represent the relation between duration d (measured in hours) and its associated depth of rainfall h (measured in millimeters). For each level-6 basin, the annual maximum value from 1998 to 2007 has been individuated for the chosen critical durations d and then the average of these values has been calculated. The results form the empirical curve (Fig. 16.4), that is interpolated by an analytical curve having the formula:

$$h = a \times d^n, \quad (1)$$

where h is the height of rainfall (mm); d , duration of rainfall (h); a and n , coefficients.

The statistical method of deriving the extreme events is done by the Gumbel distribution (Chow 1954) and permits the calculation of the recurrence interval of the extreme rainfall values, measured in years and called return time (RT). Here the return time curves has been calculated considering 2, 5 and 10 years (Fig. 16.4) and using the frequency factor (K_T) introduced by Chow (1954), function of the probability level. The formula for the K_T , expressed on the basis of Gumbel distribution is:

$$K_T = 1 - c_v \left[0.45 + \frac{\sqrt{6}}{\pi} \cdot \ln \ln \left(\frac{RT}{RT - 1} \right) \right], \quad (2)$$

where K_T is the probabilistic factor of growing; RT , the return time taken in account; c_v , the variation coefficient of the averages of the annual maxima, expressed by

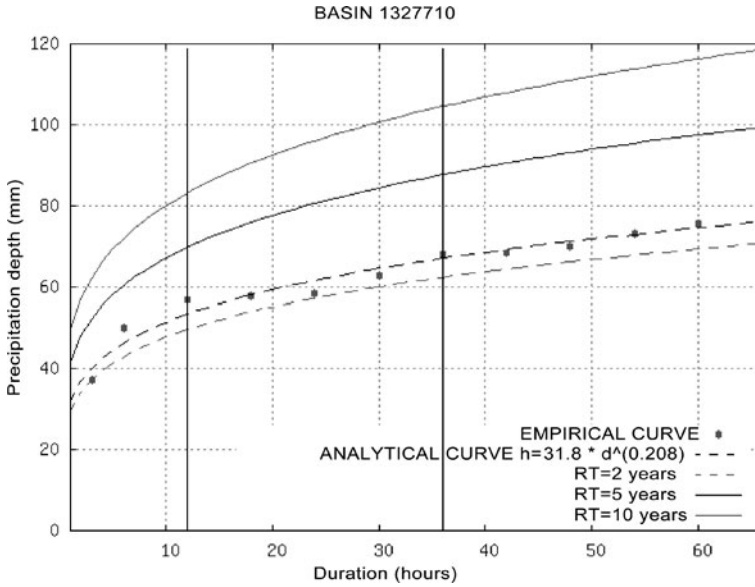


Fig. 16.4 DDF curves. The vertical lines represent the time span in which critical rainfalls had to be detected

$$c_v = \frac{\sigma_d}{\mu_d}, \tag{3}$$

where σ_d is the standard deviation of the maximum rainfall of duration d ; μ_d , the average of the maximum rainfall of duration d .

This calculation aimed at finding hydrological curves for each level-6 basin at worldwide extent has been developed using a completely automatic procedure developed in Octave© (a Matlab© porting open source software).

After having calculated these curves, the entity of each rainfall event was determined comparing cumulated rainfall to the historical ones (named analytical curve in Fig. 16.4). The whole dataset of rainfall (more than 29.000 data for a single basin) was processed following the afore mentioned procedure: the cumulated rainfalls were calculated for the specific durations and they were compared with the level-6 analytical curve. It should not forgotten that critical durations are around lag time (see Section 16.3.2), precisely chosen from $0.8 \cdot \text{lag time}$ to $2 \cdot \text{lag time}$ (delimited in Fig. 16.5 by the two vertical lines) (Alferi et al. 2008). The highest cumulated rainfall in this period was considered as a critical rainfall event (represented by a black dot, in Fig. 16.5).

The comparison between cumulated rainfall and the characteristic curve was made by using the K_T factor, considered as:

$$K_T = \frac{h_{d_1}}{ad_1^n}, \tag{4}$$

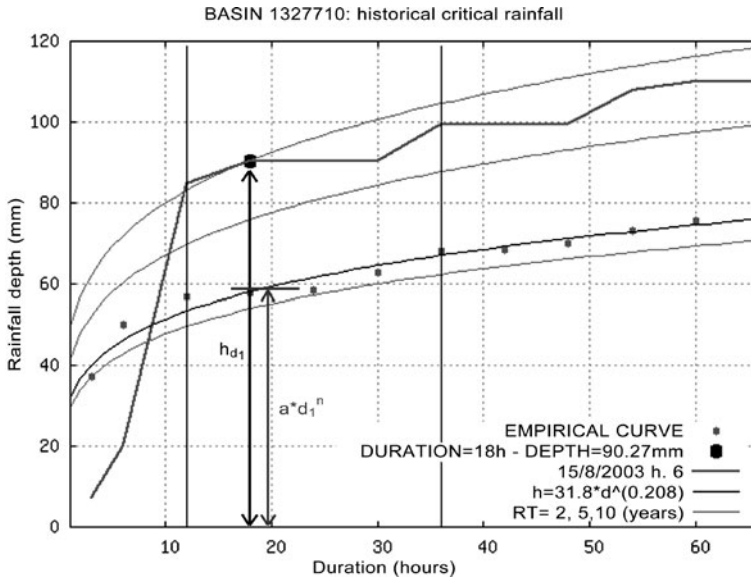


Fig. 16.5 Detection of critical rainfall, through the calculation of the K_T factor

where h_{d_1} is the cumulated rainfall for the duration d_1 (mm); d_1 , duration (h); a and n , coefficients calculated by using the formula (1).

Cumulated rainfall is considered critical when the value of K_T is higher than 0.9 (Fig. 16.6).

The list of the events found with the calculation of the K_T factor, constitute our database of historical floods and it is comparable with other archives of historical floods (e.g. Dartmouth Flood Observatory, <http://www.dartmouth.edu/~floods>).

16.3.2 Automatic Calculation of Some Drainage Basin Parameters

The calculation of river channel morphometric characteristics is important to understand the hydrologic process at basin scale.

Suitable allowance should be made in finding critical durations of rainfalls that give peak discharge; according to many authors these durations are directly related to time of concentration (t_c). It is proven that the lag from the centroid of rainfall excess to the peak of the hydrograph, known as lag time (t_l), is assumed to be $0.6t_c$ (0.6 times time of concentration) (Maidment 1992, Loukas and Quick 1996).

Time of concentration (t_c) is an idealized concept and is defined as the time taken for a drop of water falling on the most remote point of a drainage basin to reach the outlet; in other words the time of concentration is the time after commencement of rainfall excess when all portions of the drainage basin are contributing to flow at the

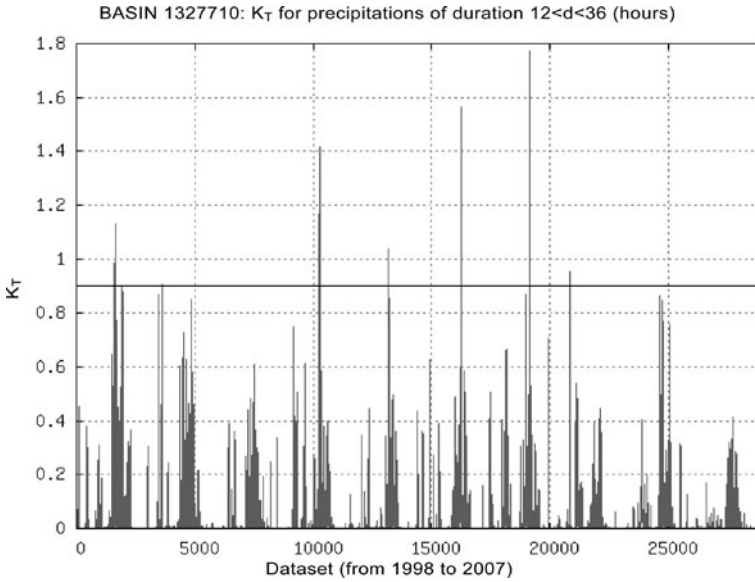


Fig. 16.6 Plot of the K_T factors found by the model, the horizontal *black line* is a threshold that shows K_T higher than 0.9

outlet (Maidment 1992). One of the empirical formula used to calculate the time of concentration (t_c) is Giandotti's (Giandotti 1934, 1940, Tonini 1959):

$$t_c = \frac{4\sqrt{A} + 1.5L}{0.8\sqrt{z_m - z_0}}, \quad (5)$$

where A is the area of the river basin (km^2); L , the length of the longest flow path (km); z_m , the medium altitude of the basin (m); z_0 , the altitude at the outlet (m).

The procedure for automatic calculation of the lag time uses HYDRO1k information (GTOPO30 DEM, HYDRO1k Stream Line layer and Drainage Basin layer). HYDRO1k provides a standard suite of geo-referenced data sets (at a resolution of 1 km) that are commonly used to process hydrologic information on a continental scale. The basis of all the data layers available in the HYDRO1k database is the hydrologically correct DEM. This DEM is projected into an equal area projection (Lambert Azimuthal Equal Area projection), the cell size is 1000 m and the radius of the sphere of influence of 6,378,137 m.

The *stream line* data layer distributed within the HYDRO 1 k data set is derived from the flow accumulation and flow direction layers. Cells with upstream drainage areas greater than 1000 km^2 are selected from the flow accumulation layer and processed through a specific function. These procedures result in a vector data layer of streamlines with each segment of stream attributed with the upstream contributing

drainage area. The vector streamlines are attributed with many fields, in particular some of them are necessary in the analysis:

Pf_type=the Pfafstetter level at which the stream segment is considered “main stream”.

Level-1 to Level-6=The Pfafstetter units in which the stream segment lie.

Strorder=Strahler stream order of the segment.

Frmup_flowlen=The upstream flowlength from the from-node. Calculated using ARC/INFO’s FLOWLENGTH function, it is the longest path from the from-node to the drainage basin divide (m).

Toup_flowlen=The upstream flowlength from the to-node (m).

Frmdn_flowlen=The downstream flowlength from the from-node. Again from ARC/INFO’s FLOWLENGTH function, it is the length from the from-node to the ocean or a terminal sink (m).

Todn_flowlen=The downstream flowlength from the to-node (m).

The *drainage basin* distributed with the HYDRO1k data set are derived using the vector streamlines along with the flow direction layer. The basins are seeded following procedures first articulated by Otto Pfafstetter and adopted for use in the HYDRO1k dataset (Verdin and Greenlee 1996). Each polygon in the basin data set has been tagged with Pfafstetter code uniquely identifying each sub-basin. The six-digit Pfafstetter code assigned to each basin carries basin linkage information. This way, basin interconnectedness can be determined by simply examining the Pfafstetter code.

The drainage basin polygons are attributed with many fields, in particular some of them are necessary in the analysis:

Level-1 to Level-6=Pfafstetter units of each polygon.

Dem_mean=Mean elevation value within the subbasin (m).

These parameters belonging to the HYDRO1k were used to calculate the time of concentration with Giandotti’s formula (5), and then the lag time, using completely automatic procedures.

The algorithm to calculate the longest flow path L was programmed in Visual basic 6. The characteristic of this procedure is to consider the river basin structure as a tree structure. It selects the level-6 basin and the consisting stream lines; then it uses the Frmdn_flowlen and Todn_flowlen fields to find the node to which corresponds only the to_node information and considers it as the outlet (Fig. 16.7). Then a recursive algorithm starts from this point, and computes the length of each river sector between two nodes up to the river basin boundary. The algorithm comes back to the outlet and compares the calculated lengths at each node to select the longest. The longest sector found is written in the drainage basin table in a specific field as the longest flow path or principle river course of the basin (Fig. 16.8).

The medium altitude of the basin z was calculated using the elevation information available in the HYDRO1k: the basin elevation characteristics were selected using

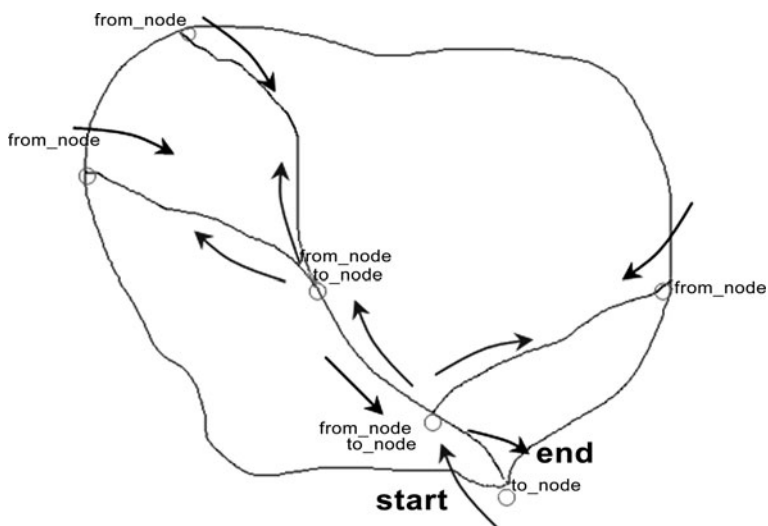


Fig. 16.7 Simplified schema of the recursive algorithm to calculate the longest flow path for a level-6 basin

ArcGIS spatial analyst tools and the average and minimum elevation related to each level-6 basin was found and reported in the drainage basin table as a new field. The area of the basin A was also calculated using ArcGIS.

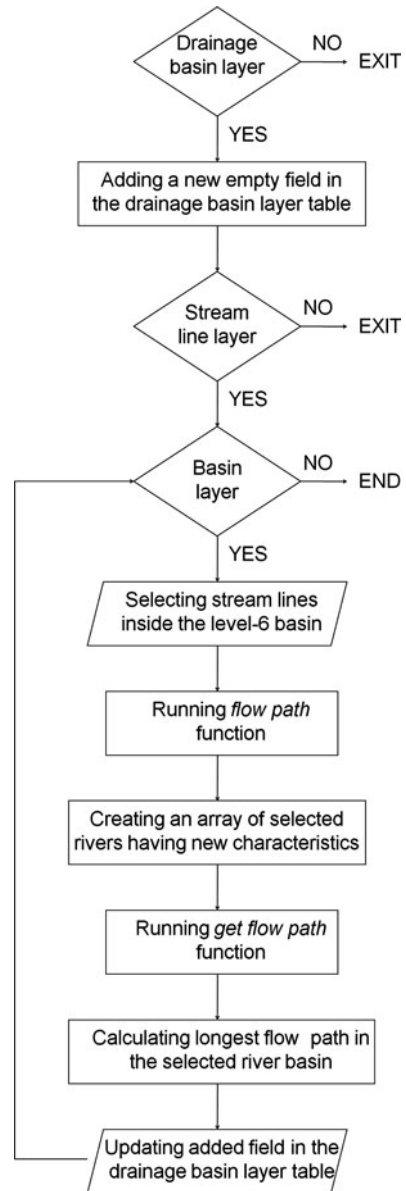
16.4 Real Time System

A Real Time System is the way to monitor rainfall severity over an area of interest in near real time. It should be able to analyze real time rainfall data to detect critical rainfalls and to present affected areas using maps. In order to achieve these requirements, ITHACA developed a Real Time System which compares the real time rainfall intensities with the curves developed by the hydrological model (see Section 16.3.1). The choice in the scientific panorama concerning real time data, excluding meteorological forecast models, is limited to the 3B42RT, an algorithm with the same coverage of 3B42, which offers rainfall data available in near real time.

16.4.1 Correction Factor Algorithm and Data Correction in Real Time

The 3B42RT belongs to the TRMM Multisatellite Precipitation Analysis TMPA (Huffman et al. 1997, 2007), a Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. The TMPA is designed to combine

Fig. 16.8 Flow chart of the Visual Basic algorithm for calculating the longest flow path for each level-6 basin



precipitation estimates from various satellite systems, as well as land surface precipitation gauge analyses when possible, with the goal of finally having a calibration traceable back to the single “best” satellite estimate.

The Real Time data are available from 2002 to today, in particular 3B42RT is represented by HQ/VAR merged data (HQ passive microwave 3B40RT, VAR

microwave-calibrated 3B41RT) with a simple replacement of the HQ with VAR when the first is missing. Estimates are posted onto the web about 9 h after observation time, although processing issues may delay or anticipate this schedule, moreover these estimates are to be considered as highly experimental, because the calibrator used for the research product (as instance 3B42), the TRMM combined instruments (TCI), is not available in real time and because for the RT system the monthly gauge adjustment is not possible.

An adjustment is important to minimize the difference between the 3B42RT and 3B42 due to differences both in calibration source and in usage of gauge data. For this purpose a statistical approach was implemented using the superimposition between 3B42 and 3B42RT available from January 2003 to December 2007. These data have been stored at geodb_id scale using an Extraction, Transformation and Loading (ETL) algorithm, written in IDL. This algorithm downloads 3B42 in .hdf format and 3B42RT in .bin format, transforms into .txt format and associates them to: geodb_id, year, month, day. It assigns -2 for negative values and deletes the values related to over sea areas. The output is a table with two values for each 3 h period from 2003 to 2007 (3B42 and 3B42RT) associated with geodb_id.

A specific algorithm was used to implement the mentioned statistical approach which considers the correlation between 3B42 and 3B42RT and classifies them in n classes on the basis of their value.

The analysis was carried out on the basis of level-6 basins which may consist up to 20000 geodb_ids, producing millions of values at worst case (one value for each 3 h period for 5 years for 20000 geodb_ids). In this analysis, the 3B42 was plotted against 3B42RT with having same subdivision for both axes. More subdivisions were allocated for the lower values as 90% of data locate in this sector (Figs. 16.9 and 16.10).

Basically the values were classified into three groups as values equal to 0, greater than 0 and less than 10, and equal or greater than 10. Then, the values greater than 0 and less than 10 were subdivided into five groups of equal interval width. The values equal or greater than 10 were also subdivided into five groups of equal interval width. This classification creates a grid structure (Fig. 16.10). For every i th class of 3B42RT the algorithm computes:

- number of 3B42 contained in the j th cell of i ,
- average of 3B42 values contained in the j th cell of i ,
- total number of 3B42 in i ,
- weighted average for every i th class, obtained dividing the number of values of the j th cell of i class by the total number of 3B42 values in i class and multiplying the result with the 3B42 average of the j th cell of i class. This will be the value to be assigned at the i th class of 3B42RT.

Since the data is available only for 5 years, it is necessary to find a method to adjust the values which may fall outside the defined classes for a particular level-6 basin, for example the values greater than the maximum recorded value in the considered period. To solve this problem a linear regression analysis was

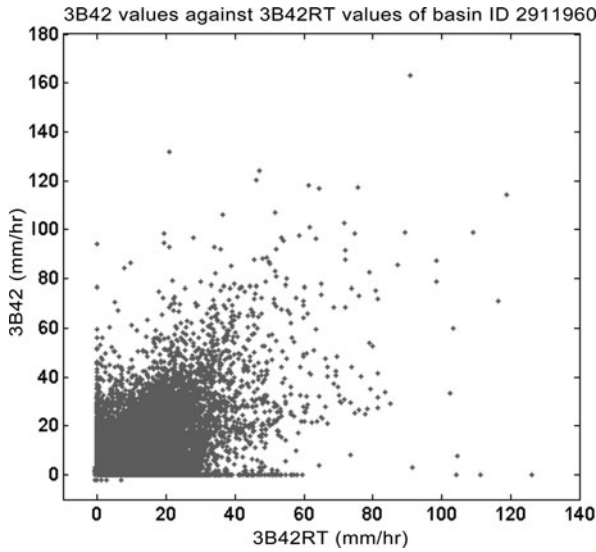


Fig. 16.9 Plot of 3B42 versus 3B42RT. It is clear how the greatest quantity of data are assembled in the bottom left corner

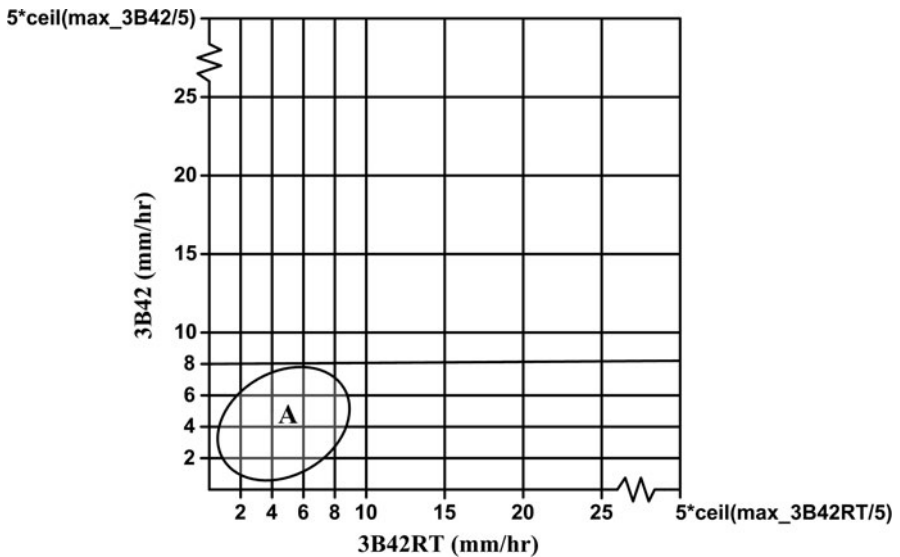


Fig. 16.10 Classification methodology of 3B42 and 3B42RT used by the ITHACA algorithm to adjust 3B42RT. The greatest quantity of data falls in the area shown as A

conducted over all data and the results were stored with correction factor results in the same file. Furthermore, the results of the regression analysis were applied for the 3B42RT classes which do not contain enough number of past data recordings.

This is a completely empirical method and its development was conditioned according to the results obtained during its construction. In fact, to validate the adjusted data they were used in the hydrological analysis to check whether it is possible to detect all the past critical events by using them as well as using 3B42 data. Through iterated tests the final version was obtained with having few variations to the primary version to develop a more generalized one which does not neglect any of past critical events. So the algorithm computes the value for each i class of 3B42RT and compares it with the value given by linear regression, and takes the greater value. Furthermore, additional weight was given for the number of counts of higher 3B42 values as they are not very common naturally, by increasing number of counts in j th cell of i th class, by multiplying a number series which increases gradually with increasing j (Fig. 16.11).

Once the methodology was tested, the entire analysis starts with the IDL ETL algorithm and ends up with the creation of correction factor files (Fig. 16.12). The IDL creates a file having both 3B42 and 3B42RT data for every 3-h period (RAIN DATA) and the combined data for a group of geodb_id(s) belonging to a level-6 basin are created (RAIN of GEODB_ID of level-6 basin), through a SQL query. Then the algorithm (CODE for the determination of the correction factor) runs to determine the correction factors. The created file for each level-6 basin contains i correction factor values and the two parameters of linear regression.

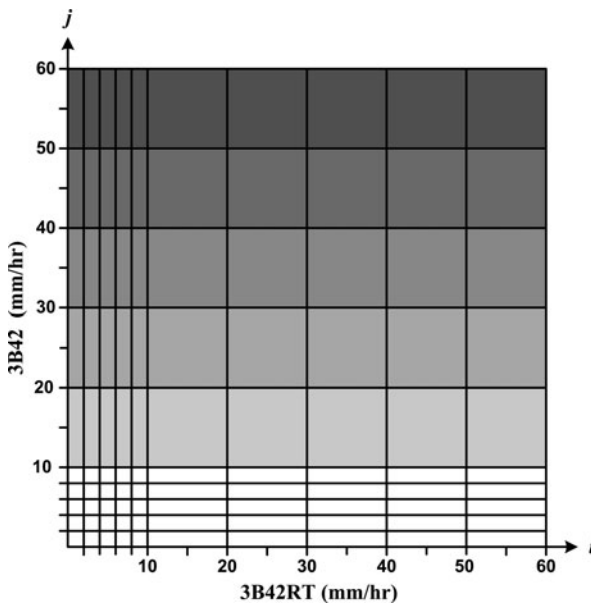


Fig. 16.11 Application of additional weight along 3B42 axis to adjust 3B42RT

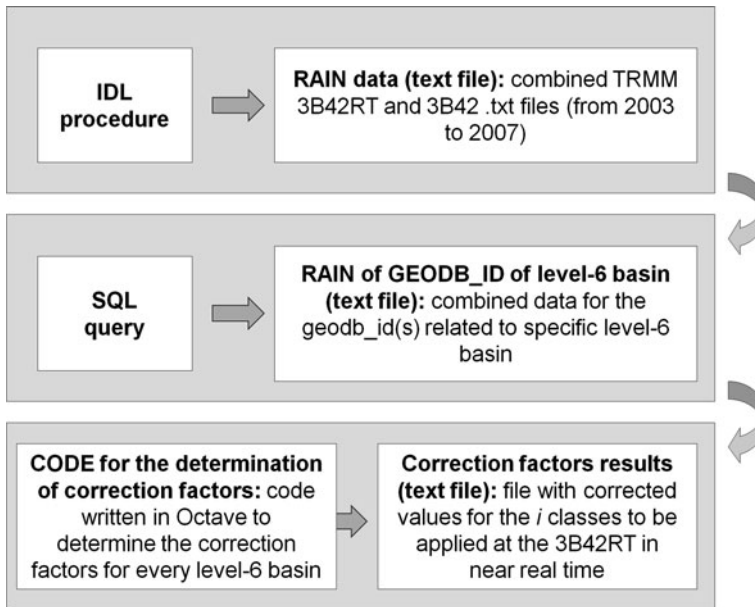


Fig. 16.12 Work flow for the correction factors determination, from the downloading of rough 3B42RT and 3B42 rainfall data to the storage of files of correction related to each specific basin

16.4.2 Detection of Critical Rainfalls in Real Time with Grid Computing System

Once the correction factor is available for all level-6 basins, it is possible to correct real time satellite rainfall data product directly when it is downloaded.

The real system is ready and it starts downloading 3B42RT from the FTP server (<ftp://trmmopen.gsfc.nasa.gov/>) and finishes saving critical events detected in the database (Fig. 16.13). It checks the server for the data availability in every 21600 s and downloads data available. Another Extraction Transformation and Loading (ETL) algorithm written in IDL is used to resample the 3B42RT cell in the geodb_id spatial resolution (see Section 16.2.2) and creates a file with the 3B42RT information. The ETL algorithm assigns zero to the negative rainfall values and does not consider data over sea areas.

Once rainfall values are downloaded and transformed into .txt format, the files having rainfall data related to a single level-6 are generated; those files are constituted by geodb_id, temporal collocation, area, 3B42RT value and an empty field that will be filled with the corrected value. This is a very important step, because within few seconds and without passing through the Oracle database, files needed by the hydrological model are created.

At this point this adjusted data can be used for the detection of critical events, by using the hydrological model above described (see Section 16.3.1).

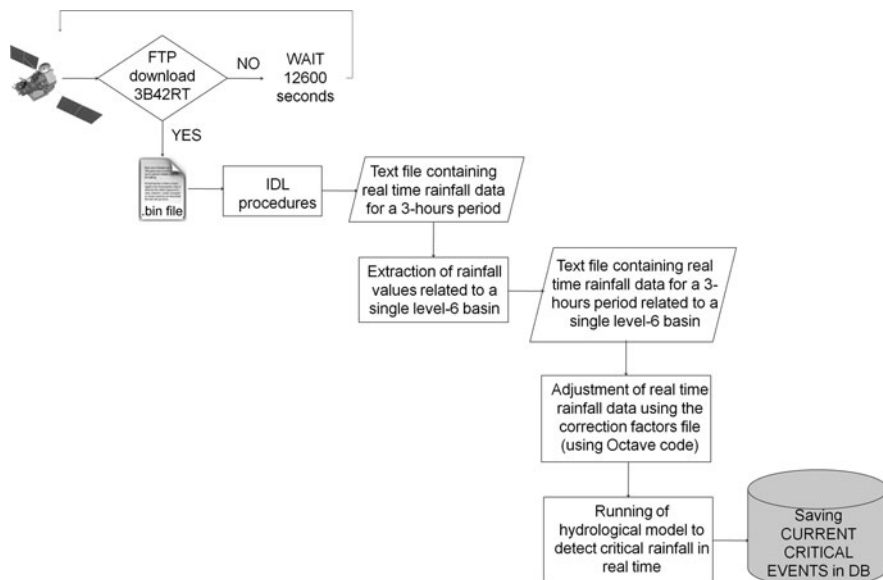


Fig. 16.13 Work flow of the Real Time System: from the downloading of 3B42RT rainfall data in real time to the detection of critical rainfall events

To run the system in real time, a grid architecture has been designed.

Grid computing (or the use of computational grids) is the combination of computer resources from multiple administrative domains applied to a common task, usually to a scientific, technical or business problem that requires a great number of computer processing cycles or the need to process large amounts of data.

A grid master (grid master node) is connected to the TRMM server and to the Oracle DBMS (Database Management System) (with historical data) and the analysis procedures are delegated to others nodes (Fig. 16.14). In particular the grid master node manages all the preliminary procedures of extraction and transformation from the primitive data, otherwise the hydrological analysis and the related detection of critical rainfalls are delegated to grid nodes. The designed network is connected to a web server to store critical events and to publish information.

The scalability of this system allows to increase or decrease in the little time the number of the nodes, depending on the data amount.

16.4.3 Results and Conclusion

Once they are triggered, alerts are automatically mapped using the informative layers extracted from WFP Spatial Data Infrastructure (SDI) (Ajmar et al. 2008b). The overall situation can be visualized on the web application, in order to offer an easy

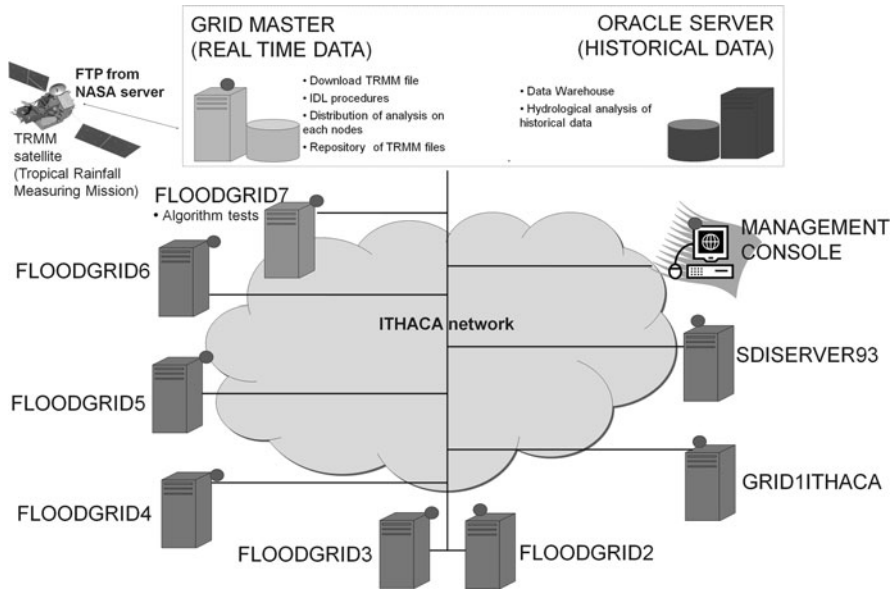


Fig. 16.14 Grid computing architecture

access to the data during the emergency also to WFP’s local offices. The web interface is developed with the most advanced javascript technologies in the geoweb field (Agosto and Dalmasso 2009).

The system shows the alerted basins with three different possible colours, according to the severity of rains; in particular, yellow colour means pre-alert (K_T parameter is near to 1), orange indicates moderate alert (K_T is between 1 and 2) and red is for heavy alert (K_T higher than 2).

At the time being, ITHACA Early Warning System for floods is under testing for evaluation of the correctness of the alerts. The system is showing good results in monitoring heavy rainfalls and giving reliable alerts for floods.

The flood that hit Western and Central Africa territories in September 2009 is an example of ITHACA’s system successful recent results. At the begin of September some of the world’s most important media involved in natural disasters monitoring announced a flood event in Burkina Faso. This piece of news was reported on September 1st by the Famine Early Warning Systems Network (FEWS NET, <http://www.fews.net>), which is a United States Agency International Development (USAID)-funded activity that collaborates with international, regional and national partners to provide timely and rigorous early warning and vulnerability information on emerging and evolving food security issues (www.fews.net). On the same day the event in Burkina Faso was given the identification code FL-2009-000172-BFA, called GLObal IDentifier number (GLIDE, www.glidenumbers.net), a globally common Unique ID code for disasters, issued every week by Emergency Events

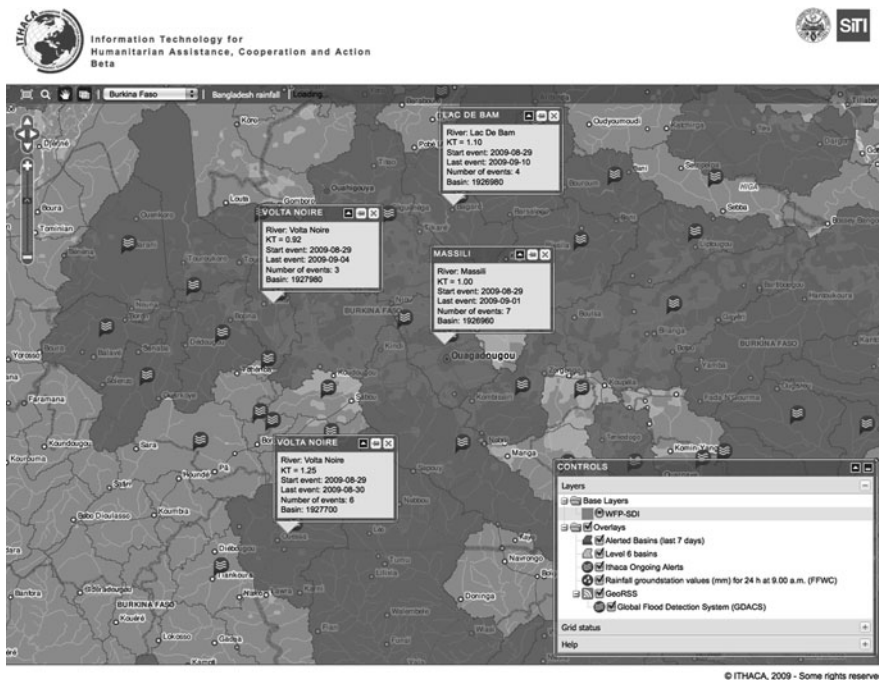


Fig. 16.15 September 2009 Burkina Faso flood: example of flood event detected by ITHACA early warning system

Database (EM-DAT, www.emdat.be) at Centre for Research on the Epidemiology of Disasters (CRED) for all new disaster events.

The same piece of news was confirmed also by ReliefWeb (www.reliefweb.int), the world’s leading on-line gateway to information (documents and maps) on humanitarian emergencies and disasters. ReliefWeb was launched in October 1996 and is administered by the UN Office for the Coordination of Humanitarian Affairs (OCHA).

ITHACA Early Warning System indicated critical rainfalls on those territories starting from August 29th (Fig. 16.15): an alert was therefore given some days in advance. Some other bordering countries, for example Niger, were also hit by heavy rainfalls in the same days; again in this case, the analysis performed by ITHACA’s system detected the critical condition in advance. The alert was confirmed during the following days and the same alerted basins turned out to be flooded areas, as showed in the map produced on September 3rd on the basis of satellite images (Fig. 16.16), through a procedure developed in ITHACA’s Early Impact Project (Ajmar et al. 2008a).

After this period of evaluation and some improvements if necessary, the Early Warning System will be available on ITHACA’s website.

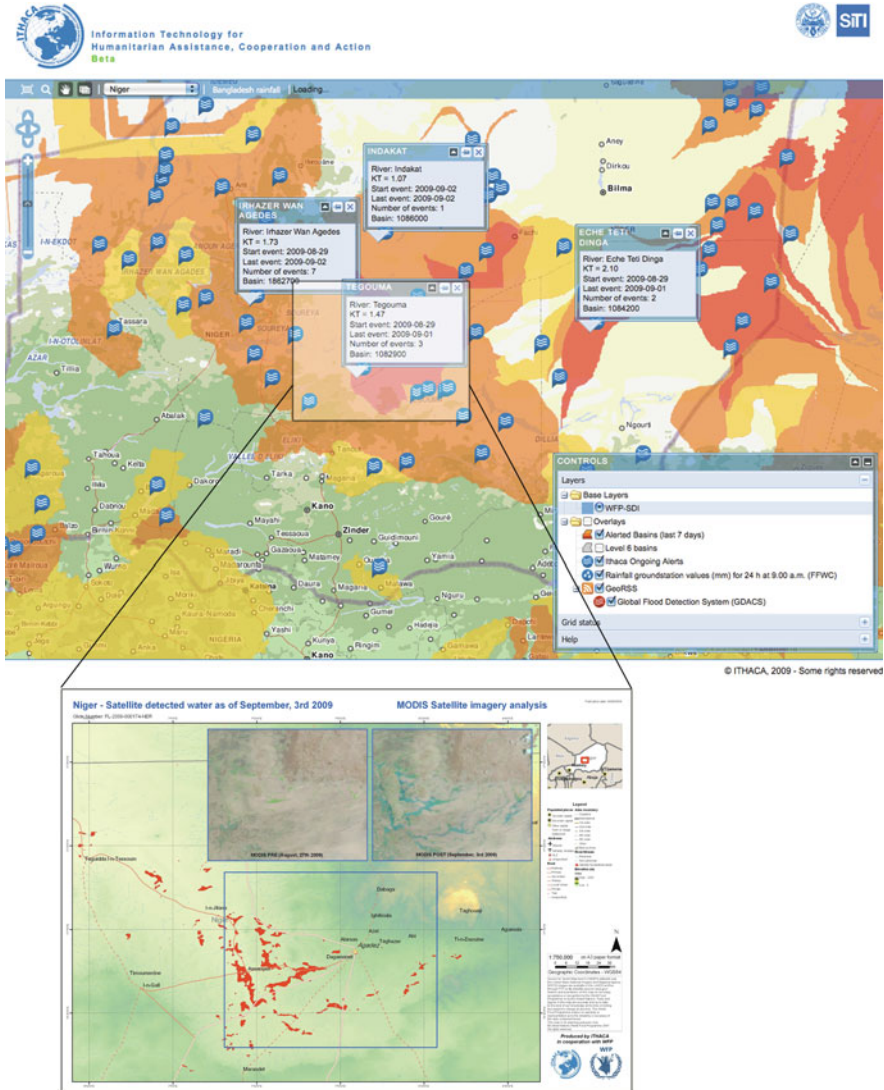


Fig. 16.16 September 2009 Niger flood: the areas alerted by ITHACA early warning system (upper image) turned out to be flooded areas in the map (lower image) produced on the basis of satellite images (Ajmar et al. 2008a)

References

Ajmar A, Boccardo P, Disabato F, Giulio Tonolo F, Perez F, Sartori G (2008a) Early impact procedures for flood events February 2007 Mozambique flood. *Italian J Remote Sens* 40(3):65–77

Ajmar A, Perez F, Terzo O (2008b) WFP Spatial Data Infrastructure (SDI) implementation in support of emergency management. *Proc XXI Congress ISPRS (Int Soc Photogrammetry Remote Sens)*, pp. 1097–1104

- Agosto E, Dalmasso S (2009) Ithaca geospatial services on a FOSS framework. Proc GI4DM19-25
- Albanese A, Disabato F, Terzo O, Vigna R, Giardino M, Perotti L (2008) A preliminary approach to flood risk mapping and flood forecasting system for the LDCs. Proc XXI Congress ISPRS (Int Soc Photogrammetry Remote Sens), pp. 1537–1542
- Alfieri L, Laio F, Claps P (2008) A simulation experiment for optimal design hyetograph selection. *Hydrol Process* 2:813–820
- Al-Sabhan W, Mulligan M, Blackburn G A (2003) A real time hydrological model for flood prediction using GIS and the WWW. *Comput Environ Urban Syst* 27:9–32
- Bates BC, Kundzewicz ZW, Wu S, Palutikof JP (eds) (2008) Climate change and water. Technical Paper of the Intergovernmental Panel on Climate Change. IPCC Secretariat, Geneva.
- Chow VT (1954) The log-probability law and its engineering applications. Proc ASCE 80:1–25
- Dartmouth Flood Observatory (DFO) <http://www.dartmouth.edu/~floods/>, Accessed 25 September 2009
- CRED/EM-DAT Emergency Events Database <http://www.emdat.be>, Accessed 25 September 2009
- FEWS NET <http://www.fews.net>, Accessed 25 September 2009
- GLIDENumber <http://www.glidenumber.net>, Accessed 25 September 2009
- Giandotti M (1934) Previsione delle piene e delle magre dei corsi d'acqua. Memorie e studi idrografici Ministero dei LL PP Servizio Idrografico Italiano 8 Pubbl 2 del S I I
- Giandotti M (1940) Previsione empirica delle piene in base alle precipitazioni meteoriche, alle caratteristiche fisiche e morfologiche dei bacini, Applicazione del metodo ad alcuni bacini dell' Appennino Ligure. Memorie e studi idrografici Ministero dei LL PP Servizio Idrografico Italiano 10 Pubbl 2 del S I I
- Golfarelli M, Rizzi S (2002) Data Warehouse, Teoria e pratica della progettazione. Mc Graw-Hill, Milano
- HYDRO1k <http://eros.usgs.gov/products/elevation/gtopo30/hydro/index.html>, Accessed 25 September 2009
- Huffman G J, Adler R F, Arkin P, Chang A, Ferraro R, Gruber A, Janowiak J, McNab A, Rudolf B, Schneider U (1997) The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull Am Meteor Soc* 78:5–20
- Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB (2007) The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. *J Hydrometeor* 8(1):38–55
- ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action) <http://www.ithacaweb.org>, Accessed 25 September 2009
- Loukas A, Quick M C (1996) Physically-based estimation of lag time for forested mountainous watersheds. *Hydrol Sci J* 41(1):1–19
- Maidment DR (1992) Handbook of hydrology. McGraw-Hill, New York
- ReliefWeb <http://www.reliefweb.int>, Accessed 25 September 2009
- Schanze J (2009) Editorial J Flood Risk Management 2(3):149–150
- SiTI (Istituto Superiore sui Sistemi Territoriali per l'Innovazione) <http://www.siti.polito.it>, Accessed 25 September 2009
- Smithson P, Addison K, Atkinson K (2002) Fundamentals of the physical environment, 3rd edn. Routledge Taylor & Francis Group, London
- Tonini D (1959) Elementi di Idrografia e Idrologia. Libreria Universitaria, Venezia
- TRMM 3B42 DATA <http://trmm.gsfc.nasa.gov/3b42.html>, Accessed 25 September 2009
- TRMM 3B42RT data downloading <ftp://trmmopen.gsfc.nasa.gov>, Accessed 25 September 2009
- Verdin KL, Greenlee SK (1996) Development of continental scale digital elevation models and extraction of hydrographic features. Proceedings of the 3rd International Conference/Workshop on Integrating GIS and Environmental Modeling, pp. 1–26
- WFP (World Food Programme). <http://www.wfp.org>, Accessed 25 September 2009

Chapter 17

L-Band and C-Band Combined Interferometric Monitoring of the Wenchuan Earthquake

Kui Zhang, Alex Hay-Man Ng, Linlin Ge, Yusen Dong, and Chris Rizos

Abstract Satellite differential radar interferometry (DInSAR) has played a pivotal role in observing the ground deformation caused by seismic events in recent years. The Ms8.0 Wenchuan Earthquake occurred on 12 May 2008 have caused tremendous losses in lives and properties of the people in affected areas. In this paper, the surface co-seismic deformation field of the quake was mapped by the two-pass DInSAR technique. Several pairs of long strip PALSAR images and one pair of ASAR ScanSAR images were processed. The ascending (PALSAR) and descending (ASAR) results were combined to derive vertical and easting components of the co-seismic deformation field. According to the comparison between the vertical and easting deformation maps and the GPS observations, it is believed that the accuracy of the DInSAR measurement of the co-seismic deformation is affected by the atmospheric artifacts. Due to the reliable ASAR ScanSAR interferometric signals only exist in the area located at the southeast of the fault zone, more descending DInSAR results are required to obtain the vertical and easting co-seismic deformation field in the northeast.

17.1 Introduction

On 12 May 2008, the Ms 8.0 Wenchuan Earthquake occurred on the intersection of the Sichuan Basin and the Qinghai-Tibet Plateau (CEA 2008), which was one of the most destructive earthquakes in China since the 1976 Tangshan Earthquake. The economic loss caused by this quake is estimated to be 75.6 billion US dollars (epochtimes.com 2008). As of 21 July 2008, 69197 people are confirmed dead and

K. Zhang (✉)

Cooperative Research Centre for Spatial Information and School of Surveying and Spatial Information Systems, The University of New South Wales, Kensington, Sydney, NSW 2052, Australia
e-mail: kubic.zhang@gmail.com

18222 listed as missing (sina.com.cn 2008). Thousands of aftershocks, some with magnitudes greater than Ms 6.0, continued to strike the region even months after the main shock, causing new losses in lives and properties of the people in affected areas.

The Phased Array type L-band Synthetic Aperture Radar (PALSAR) and Advanced Synthetic Aperture Radar (ASAR) are two of the most widely used spaceborne SAR sensors in interferometric applications.

PALSAR is the SAR sensor installed on Japanese Advanced Land Observing Satellite (ALOS; Rosenqvist et al. 2007). Being an upgraded sensor of the one on JERS-1 satellite, the L-band PALSAR is able to acquire high quality SAR images. With 23.6 cm wavelength, PALSAR has made great contributions to the study of earth observation.

Operating in C-band, ASAR is one of the most important instruments of the environmental satellite (ENVISAT) which is the subsequent mission of ERS-1 and ERS-2 (Demos et al. 2000). ASAR is able to operate in five modes for different purposes, which are Image Mode (IM), Wave Mode (WV), Alternating Polarisation Mode (AP), Global Monitoring Mode (GM) and Wide Swath mode (WS). In WS mode (ScanSAR), ASAR is able to acquire data in five swaths so that the spatial coverage can be extended to over 400 km by 400 km.

DInSAR is a powerful tool to observe the ground surface deformation, first noted by Gabriel et al. (1989). It has been applied to a wide range of applications in the past 10 years and more (Goldstein et al. 1993; Lanari et al. 1998; Ng et al. 2009). This technique makes use of multiple coherent SAR images acquired over the same area to explore the difference between them. One of the most important products generated by DInSAR is the differential interferogram. A high quality differential interferogram is able to provide continuous observation with large spatial coverage. Comparing with other existing techniques, this fascinating advantage enables DInSAR to be extremely competitive in the monitoring of ground surface change. By using images acquired at the same location with different heading directions, DInSAR has been extended to determine 2D or 3D ground deformation in recent years (Ferretti et al. 2007; Fialko et al. 2001; Wright et al. 2004; Yun et al. 2006). With the launch of numerous satellite-based SAR sensors, DInSAR has been repeatedly utilised to map the ground surface deformation caused by seismic events (Ge et al. 2009; Jónsson et al. 2002; Massonnet and Feigl 1995). These results have exerted profound influence on the research of seismology.

After the earthquake, DInSAR was carried out to map the co-seismic deformation by using PALSAR data provided by the Japanese Earth Remote Sensing Data Archive Center (ERSDAC) and ASAR ScanSAR data provided by European Space Agency (ESA). Due to the coverage constraint of standard PALSAR images, six pairs of long strip PALSAR products were used to ensure sufficient spatial coverage for mapping the co-seismic deformation field. The grey rectangles in Fig. 17.1 represent the areas covered by each PALSAR interferometric pair. On the other hand, the ASAR ScanSAR image pair is able to cover the most of affected areas. Its coverage is indicated by the red rectangle in Fig. 17.1. All SAR image pairs and their key parameters used in this study are listed in Table 17.1.

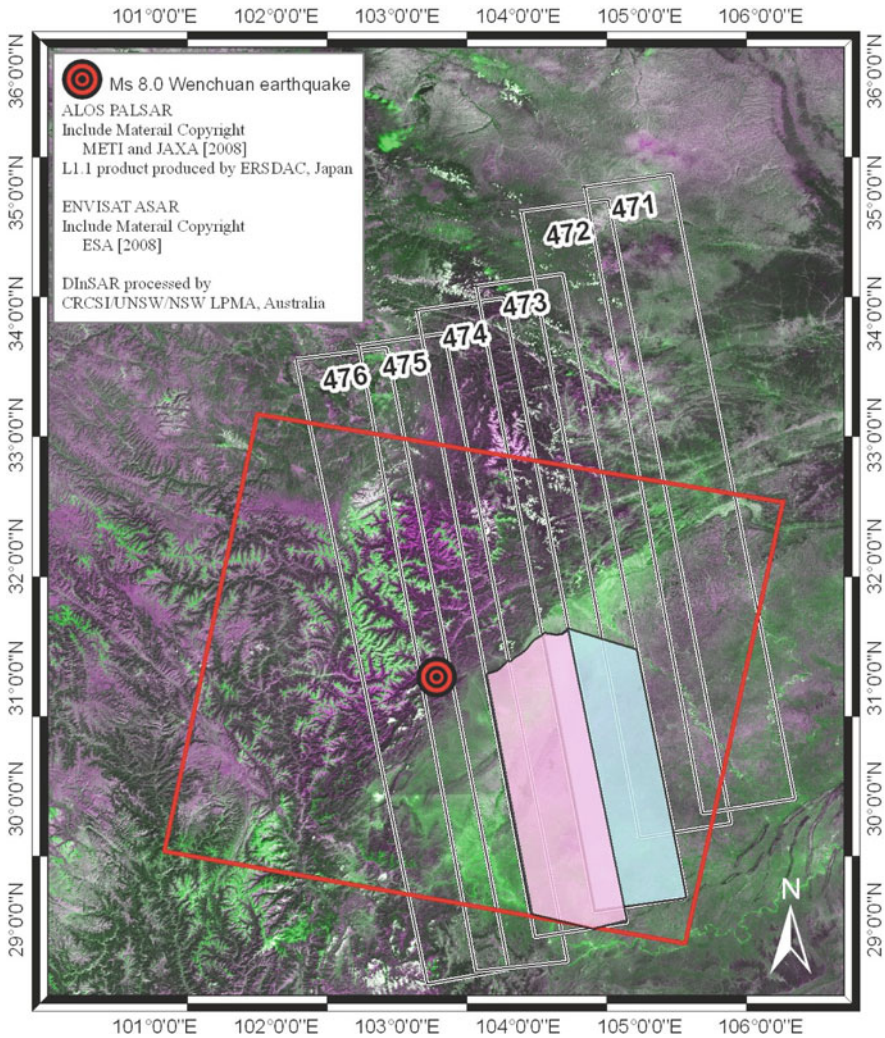


Fig. 17.1 The study area represented by Landsat-7 images. The epicentre is indicated by the large red circle. The grey and red rectangles represent the area covered by L-band PALSAR and C-band ASAR ScanSAR result, respectively. The polygons coloured by blue and pink indicate the areas where the vertical and easting deformation maps are generated

As discussed later, the quality of each PALSAR interferometric product is generally good. To some extent, the distribution of the co-seismic deformation can be observed from the L-band PALSAR DInSAR results. Unfortunately, most part of the C-band ASAR differential interferogram are severely influenced by the decorrelation. The reliable interferometric signals only exist in the areas with flat terrain. Due to the PALSAR and ASAR images used in this study were acquired on ascending and descending orbits respectively, the vertical and easting deformation maps

Table 17.1 PALSAR and ASAR image pairs used

Sensor	Path/track	Acquisition date	Perpendicular baseline (m)	Temporal baseline (d)	Look angle	Headings
ASAR	247	25 January 2008 13 June 2008	89.8	140	30.8–42.7	Descending
PALSAR	471	29 February 2008 31 May 2008	123.1	92	34.3	Ascending
PALSAR	472	28 January 2007 17 June 2008	237.5	506	34.3	Ascending
PALSAR	473	17 February 2008 19 May 2008	248.7	92	34.3	Ascending
PALSAR	474	05 March 2008 05 June 2008	293.6	92	34.3	Ascending
PALSAR	475	20 June 2007 22 June 2008	–41.6	368	34.3	Ascending
PALSAR	476	08 April 2008 24 May 2008	–198.1	46	34.3	Ascending

in the reliable overlapping regions between PALSAR and ASAR interferograms are calculated in this study. By comparing with the GPS observations, the accuracy of the resulting vertical and easting components of co-seismic deformation field is given.

In this paper, the general two-pass DInSAR processing strategy and the least square method for deriving the vertical and easting component of the displacement are briefly reviewed in Section 17.2. Then, the DInSAR results derived from PALSAR and ASAR images are presented. The vertical and easting deformation maps derived from DInSAR results are generated and validated with the GPS observations.

17.2 Methodology

In this section, the two-pass DInSAR approach is firstly reviewed. Then, how to estimate the vertical and easting components of the displacement by using the ascending and descending DInSAR results is discussed.

17.2.1 Two-Pass DInSAR Approach

There are several methods of generating a differential interferogram, such as the two-pass method, three-pass method, and four-pass method. The two-pass DInSAR approach was used to generate the co-seismic deformation map in this study. The reason why the three-pass and four-pass DInSAR methods were not used for the deformation analysis is based on the following considerations. Firstly, the atmospheric disturbances might seriously decrease the accuracy of space-borne

interferometric measurements. If the atmospheric delay errors exist in the “topographic pair”, the quality of the differential interferogram could not be guaranteed. Furthermore, the area affected by the main shock, as well as aftershock sequence, is mainly dominated by mountains with steep slopes. The accuracy of the interferometric measurement over this area is strongly influenced by the serious foreshortening and layover effects. In the worst case, the interferometric phase could be totally uncorrelated. It is extremely difficult to correctly obtain the topographic information from the “topographic pair” in three-pass and four-pass methods.

The two-pass DInSAR technique was firstly demonstrated in 1993 (Massonnet et al. 1993). The following is a brief review of the steps involved in two-pass DInSAR processing as illustrated in Fig. 17.2. Two radar images (referred to as ‘master image’ and ‘slave image’) are required. To begin with, the slave image is carefully coregistered to the master image. Subsequently, the raw interferogram is generated by complex conjugated multiplication of the interferometric image pair data. The raw interferogram mainly consists of the following components: (1) the

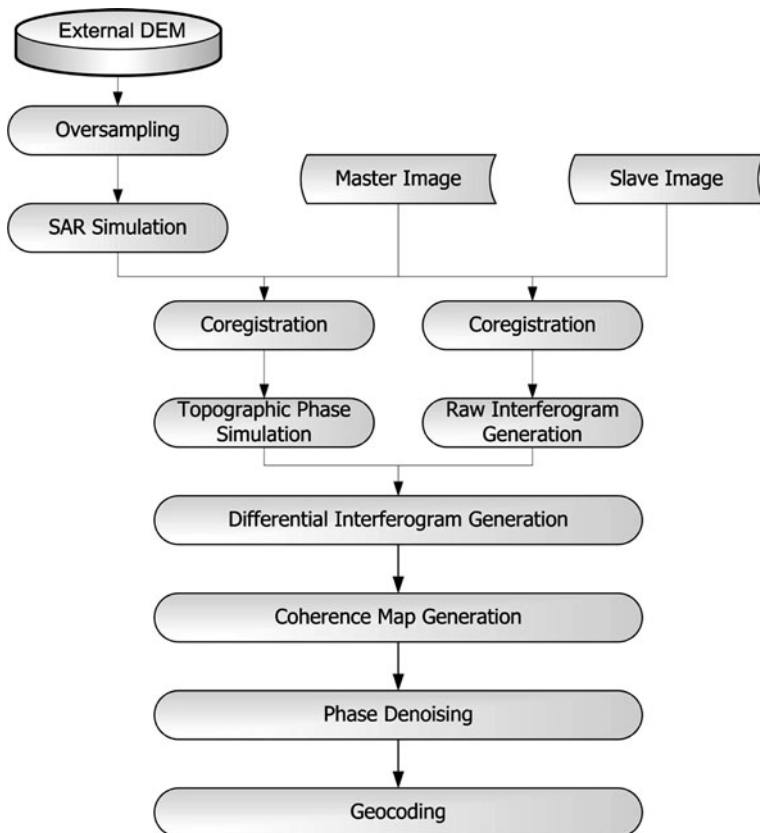


Fig. 17.2 The simplified two-pass DInSAR processing chain

phase induced by the pre-defined ellipsoid reference surface, (2) the phase induced by topography, (3) the phase induced by the deformation that occurred between master and slave image acquisitions and (4) the phase contributed by atmospheric disturbances.

The phase caused by the pre-defined ellipsoid reference surface can be directly removed from the raw interferogram by using the orbit information of both master and slave acquisitions. By utilising an external DEM, the phase contributed by the topography can be simulated and hence the differential interferogram is obtained. Several filter techniques (Fornaro and Guarnieri 2002; Goldstein and Werner 1998; Lee et al. 1998) can be used to denoise the differential interferogram. Due to the phase values observed from the interferogram are known only modulo 2π rad, a phase unwrapping step is required to recover the wrapped interferometric phase to the absolute phase with respect to the ground deformation. Finally, the unwrapped phase array is geocoded to a reference coordinate system (e.g. WGS84).

17.2.2 Estimation of Horizontal and Vertical Components

The displacement measured by both DInSAR techniques is in the line-of-sight (LOS) direction of the radar. The displacement vector along the LOS acquired by SAR satellite is a composite of vertical, easting and northing displacement components. The full 3D ground displacement vector can be resolved using three or more independent LOS measurements, unless additional assumptions on the deformation exist (Ferretti et al. 2007). The measured displacement vector along the LOS can be expressed as (Fialko et al. 2001):

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \cos(\alpha) & \sin(\theta) \sin(\alpha) \end{bmatrix} \begin{bmatrix} D_U \\ D_E \\ D_N \end{bmatrix} = \Delta R,$$

where θ is the incidence angle at the reflection point; α , azimuth of the satellite heading vector (positive clockwise from North); D_U , displacement in vertical direction; D_E , displacement in easting direction; D_N , displacement in northing direction; ΔR , ground displacement in the LOS direction between two acquisitions.

However, due to the lack of SAR data acquired from different viewing angles and orbit headings over a similar period and the poor sensitivity in measuring deformation in the north-south direction, the displacement in vertical and easting directions are calculated by supposing that the north-south component is null. In fact, azimuth of the satellite heading vector (positive clockwise from the north) for ascending (PALSAR) and descending (ASAR) orbits are 9.5° and 190.3° , respectively. It is possible to approximate these values to 0° and 180° , similar to Yun et al. (2006). With this approximation, the vertical and easting displacement components can be calculated according to the following equations:

$$\left[\cos(\theta) \mp \sin(\theta) \cos(\alpha) \right] \begin{bmatrix} D_U \\ D_E \end{bmatrix} = \Delta R,$$

where the negative sign is for PALSAR data from ascending orbit and the positive sign is for ASAR data from descending orbit.

17.3 DInSAR Results

17.3.1 PALSAR

The final DInSAR results of PALSAR are illustrated in Fig. 17.3. To express the co-seismic deformation field more distinctly, the final unwrapped results are rewrapped modulo 4π rad (two cycles). Each fringe in the resulting products represents 23.6 cm of movement in the radar line-of-sight (LOS) direction which is indicated by orange arrow in Fig. 17.3. The co-seismic deformation field of the Wenchuan Earthquake can be generally observed. The quality of the results in the northwest of the fault zone is lower than in the flat regions in the southeast. This is presumably due to: (1) extensive vegetation coverage over the mountainous regions, (2) the earthquake caused more intensive deformations in the northwest of the seismic fault, and (3) the decorrelations caused by the severe foreshortening and layover effects over the region with rough terrain.

The three arc-second resolution DEM from the Space Shuttle Topography Mission (SRTM) was used to estimate the topographic contributions to the interferometric phase. In general, the absolute vertical accuracy of the SRTM DEM is better than 9 m (Farr et al. 2007). However, they have to be oversampled to satisfy the processing requirement due to their low resolution. Inevitably errors will be introduced during the oversampling, and hence these errors will propagate into the deformation results. In particular, the region located at the northwest of the fault zone is dominated by mountains with steep slopes, where the topographic contribution to the interferometric phase could not be completely removed.

To ensure the reliability of the phase unwrapping, the areas along the fault zone were masked in this study, as can be seen from Fig. 17.3. The correlation of the interferometric phase signals in these areas is totally lost. The decorrelation is most likely due to one or more of the following reasons. Firstly, the phase difference between two adjacent resolution elements may exceed one cycle. Secondly, large ground movement can lead to misregistration of image pairs. Lastly, the surface of scatterers has changed at the wavelength scale because of the earthquake (Zebker et al. 1994). It must be noted that each differential interferogram of path 472–475 is separated into two parts by the mask areas. Therefore, the phase unwrapping has to be carried out for each part separately.

It is evident that there are some abnormal fringes in the regions with shadow mask in Fig. 17.3. These regions are far away from the fault zone and hence are expected to be less affected by the quake. The interferometric phase signals far from the co-seismic deformation field are most likely to be contaminated by the atmospheric disturbances.

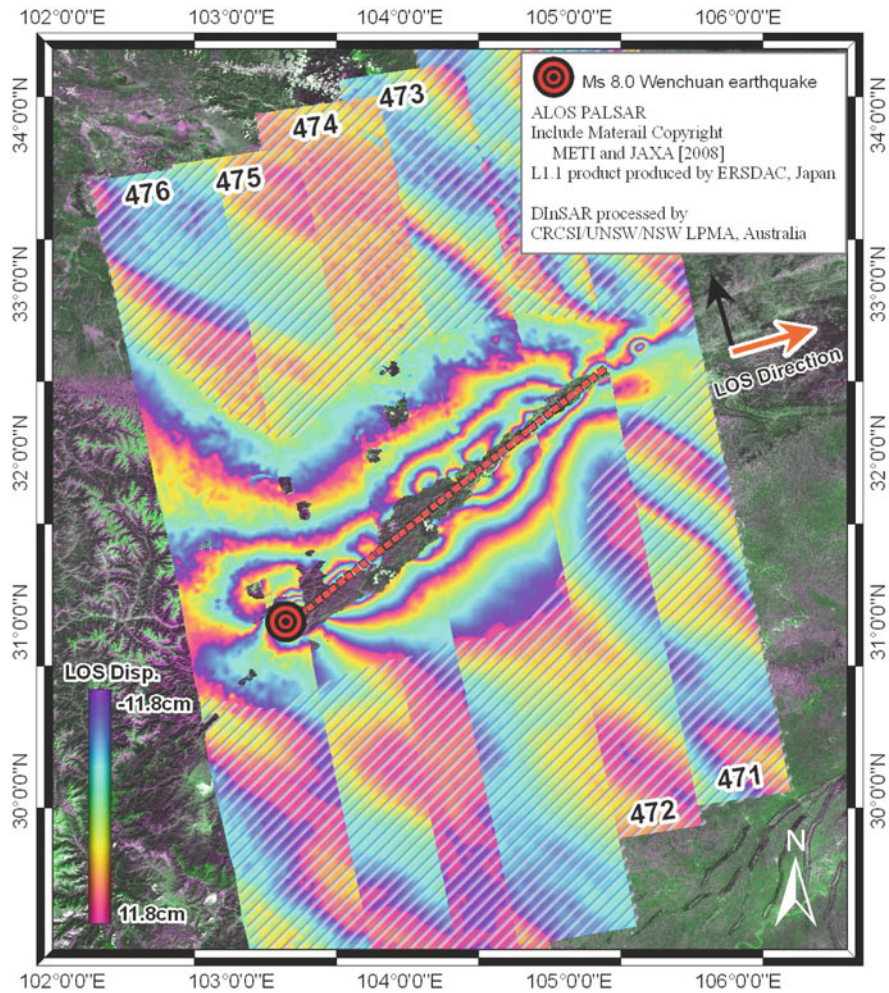


Fig. 17.3 Final PALSAR DInSAR results overlaid on Landsat-7 images. To express the co-seismic deformation filed more distinctly, the final unwrapped results were rewrapped modulo 4π rad (two cycles). The interferometric signals in the area with shadow mask are expected to be dominated by the atmospheric disturbances. The *black* and *orange* arrows represent the satellite flight pass and radar LOS direction, respectively

17.3.2 ASAR

The ASAR ScanSAR differential interferogram is represented by the greyscale layer in Fig. 17.4. Apart from the area with flat terrain, the quality of the differential interferogram is generally quite poor. Especially for the areas with rugged terrain and heavy vegetation, the co-seismic deformation signals cannot be resolved by the C-band ASAR interferogram. On the other hand, the quality of the L-band PALSAR

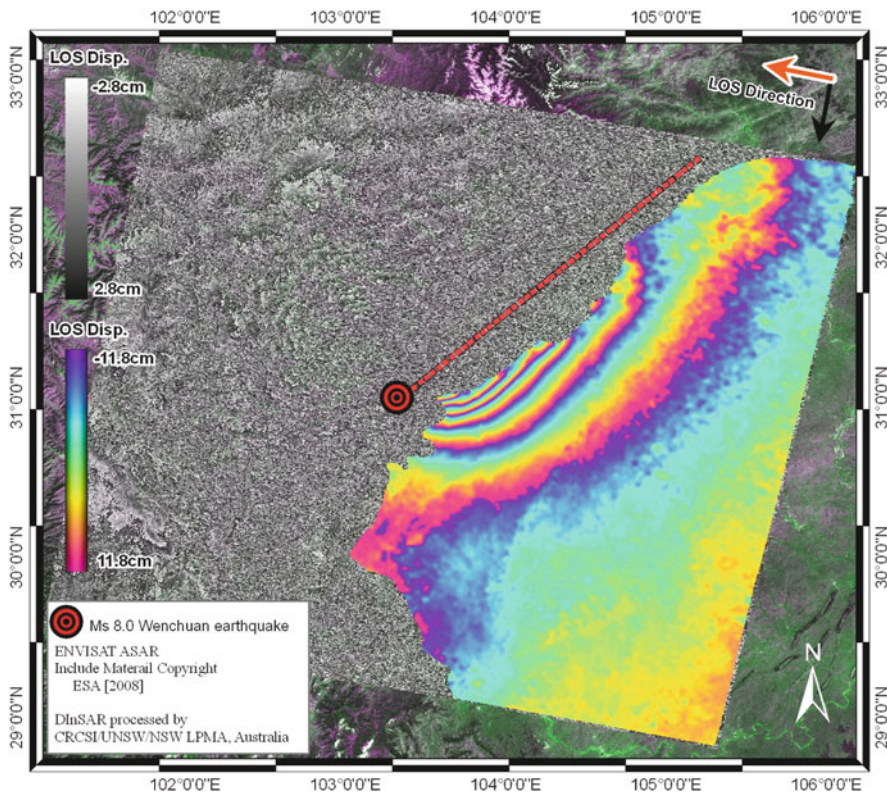


Fig. 17.4 The ASAR ScanSAR DInSAR result. The *greyscale* layer represents the differential interferogram. The area with high coherence was unwrapped. The unwrapped result was rewrapped 8.4π rad (4.2 cycles). The rewrapped result is demonstrated by the coloured layer overlapped on the differential interferogram

interferograms is generally good over these areas. This is mainly due to the L-band PALSAR sensor has a longer wavelength than the C-band ASAR sensor. Signals with longer wavelength have a better capability of penetrating heavy vegetation, which leads to less temporal decorrelation (Takeuchi and Yamada 2002).

In this study, the area with high coherence was unwrapped. In order to be more easily compared with PALSAR results, the unwrapped result was rewrapped 8.4π rad (4.2 cycles). The rewrapped result is demonstrated by the coloured layer overlaid on the differential interferogram. Same as above, each fringe in the coloured layer implies 23.6 cm of movement in the ASAR LOS direction. It can be observed from Figs. 17.3 and 17.4 that in the areas located at the southeast of the fault zone, the ground deformation along the LOS direction of ASAR is greater than that of PALSAR. In addition, the co-seismic deformation field derived from the rewrapped result of ASAR appears more consistent compared to the PALSAR results.

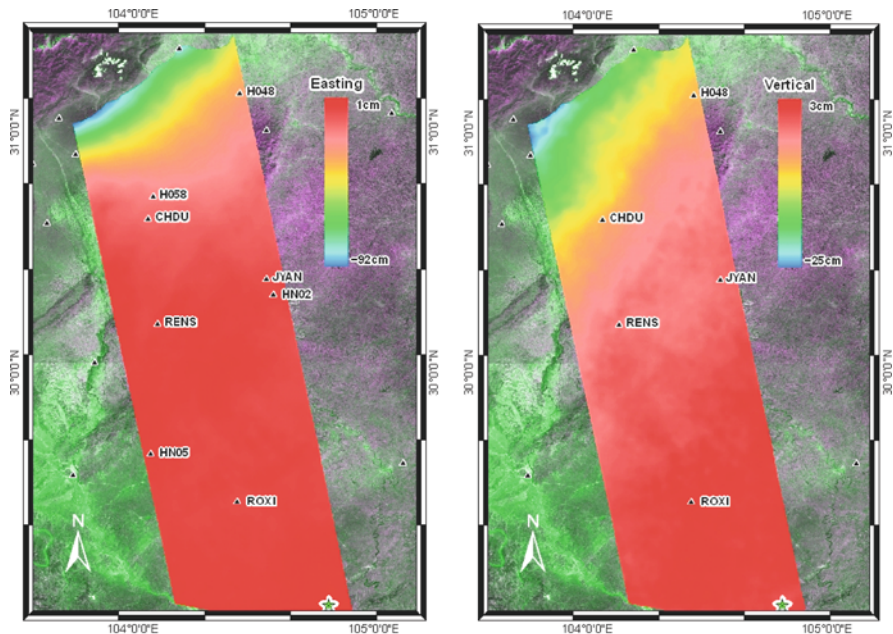
17.4 Combined Observation

In order to properly combine the LOS estimations derived from PALSAR and ASAR, both the ascending and descending interferograms must have sufficient quality in the selected region. The region which met such requirement was identified by the ASAR result. The area which preserves good coherence in the ASAR result is located at the southeast part as highlighted in Fig. 17.5. The corresponding PALSAR data which covered the same area, i.e. south parts of Paths 473 and Path 474 (as shown in Fig. 17.1), were selected. The vertical and easting displacement maps were derived by combining the ASAR and PALSAR interferograms based on the assumption that the northing component was null. Two independent combinations, ASAR-PALSAR Path 473 and ASAR-PALSAR Path 474, were used to derive the vertical and easting displacement maps. In fact the GPS observation data (CMONOC 2008) suggested that the northing displacements were generally small in the region, therefore it was expected that the northing displacements should not have large effect to the vertical and easting displacements calculated. The resulting vertical and easting deformation maps for ASAR-PALSAR Path 473 and ASAR-PALSAR Path 474 are shown in Fig. 17.5. The results suggested that the horizontal displacements due to the quake were greater than the vertical displacements in the region of observation. The easting displacement maps in Fig. 17.5 show that the relative movement of south part of Path 473 and Path 474 is towards the epicentre with a peak of 92 cm. The vertical displacement maps show that the amount of subsidence increases toward the fault. The peak vertical displacement reaches -25 cm in the selected region. It can be also observed from Fig. 17.5 that there are some abnormal deformation patterns in the areas far away from the fault zone. This is expected to be caused by atmospheric artifacts.

17.5 Validation

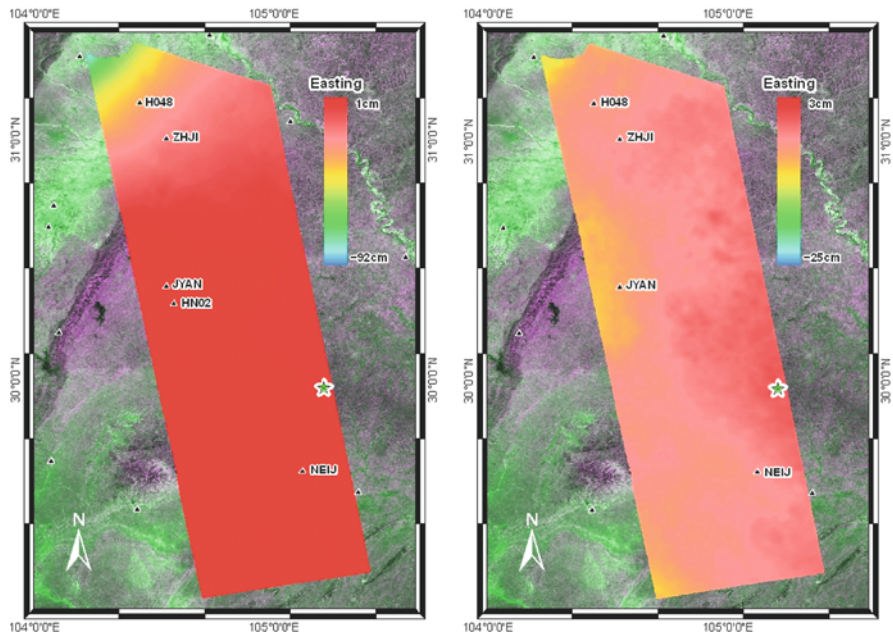
To validate the results obtained, the resulting deformation in both directions were compared against GPS data (CMONOC 2008). In this period, GPS observation data provided 10 measurement points in horizontal direction and 7 measurement points in vertical direction, located within the region covered by the vertical and easting displacement maps. The location of the GPS observations corresponding to each displacement maps are shown in Fig. 17.5. Since DInSAR measures relative phase changes, a double difference comparison between the GPS observations and the corresponding displacement component derived from the DInSAR results has been carried out.

By comparing the easting displacement maps with the GPS observation data, an acceptable correlation is observed between the DInSAR-measured and the GPS-measured easting displacements (Fig. 17.6). It is observed that the double difference in near field (points that are closed to each other) is smaller than the double difference in far field (points that are far away from each other). The double difference



(a) ASAR – PALSAR path 474 (Easting)

(b) ASAR – PALSAR path 474 (Vertical)



(a) ASAR – PALSAR path 473 (Easting)

(b) ASAR–PALSAR path 473 (Vertical)

Fig. 17.5 The vertical and easting deformation maps derived by the proposed method. The reference points for each deformation map are illustrated by *green stars*. The *triangles* represent the locations of the GPS observations (CMONOC 2008) used in the study

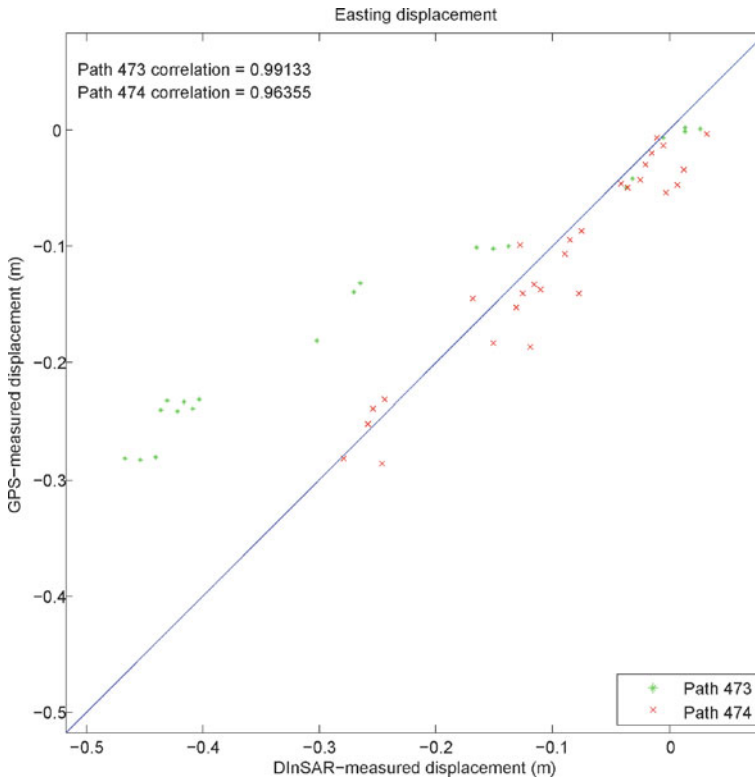


Fig. 17.6. Comparison between the DInSAR-measured and GPS measured easting displacements

observations for the points ROXI and NEJI are much higher than others. The average absolute difference between GPS-measured and DInSAR-measured easting displacements in Path 474 is 2.4 cm with a standard deviation of 1.9 cm and in Path 473 is 10.6 cm with a standard deviation of 7.5 cm. However, if only double difference observations in near field are considered, the average absolute difference for Path 473 drops to 1.8 cm and the standard deviation to 1.8 cm. This is because the effect of atmospheric signal is not obvious between nearby points and hence the deformation signal is dominated. The atmospheric signal, which is not modeled, is expected to increase for points further away from each other. The same correlation cannot be observed when comparing the vertical displacement maps with the GPS observation data, which is mainly because the amount of available GPS observation points are smaller and the points are far apart (Fig. 17.7). The average absolute difference between GPS-measured and DInSAR-measured vertical displacements in Path 474 is 4.6 cm with a standard deviation of 3.3 cm and in Path 473 is 2.0 cm with a standard deviation of 1.3 cm. The accuracy of retrieved vertical displacement results was lower than the easting displacement in this study. This is because the dispersion in the easting components of the sensitivity vectors is higher compared to the vertical components; therefore the precision of the easting displacement calculated is better.

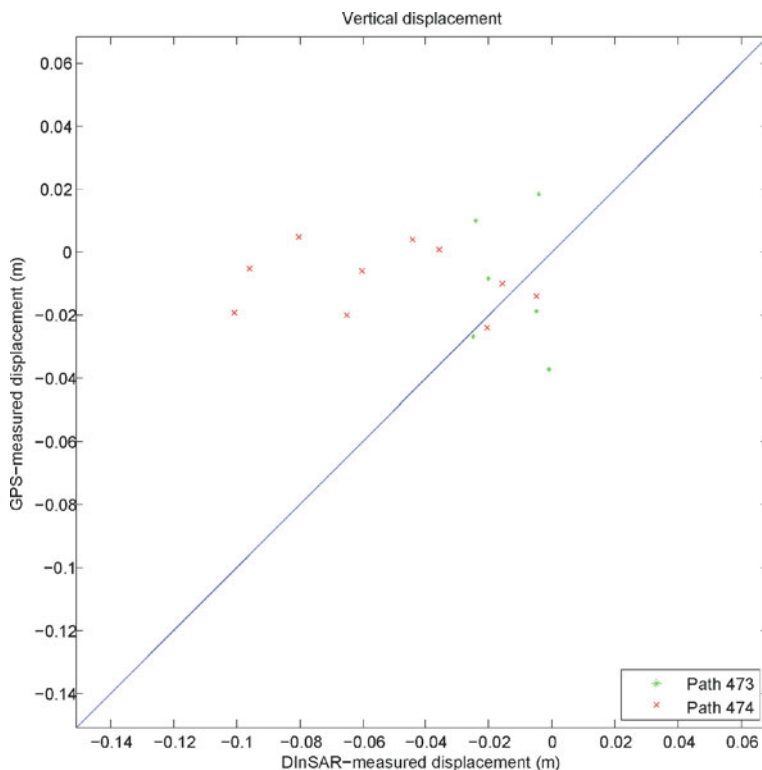


Fig. 17.7 Comparison between the DInSAR-measured and GPS measured vertical displacements

17.6 Conclusion

The capability of ALOS PALSAR and ENVISAT ASAR ScanSAR for co-seismic deformation mapping has been demonstrated in this study. While the ENVISAT ASAR ScanSAR has a far larger spatial coverage than ALOS PALSAR, most part of the ASAR ScanSAR differential interferogram are severely decorrelated. The reason is that the wavelength of C-band microwave is shorter than L-band, which causes the ASAR sensor more sensitive than PALSAR to the vegetation area. As a result, for co-seismic deformation mapping, PALSAR DInSAR is more robust in the presence of vegetation ground cover, and also more suitable for areas with high rate of ground deformation.

In the area where the interferometric signals are reliable, the results of both PALSAR and ASAR are combined to generate the vertical and easting components of the co-seismic deformation. It can be seen from the retrieved deformation map that the vertical displacements are smaller than the horizontal displacements in the region of observation. To confirm the reliability of the results, the resulting vertical and easting deformation maps were validated with the GPS ground

surveying data. The resulting easting displacements appear more consistent with the GPS observations than the vertical displacements. According to the double difference comparison, the major factor that affects the DInSAR measurement of the co-seismic deformation of the quake is believed to be atmospheric artifacts. If the interferometric phase signals contributed by the atmospheric disturbances could be eliminated, the quality of the resulting deformation maps is able to be further improved. The valid phase signals in ASAR ScanSAR differential interferograms only covered the regions located at the southeast of the fault zone. To generate vertical and easting components of the co-seismic deformation field in the northwest, more SAR images acquired in descending orbits are required.

Acknowledgments This research work has been supported by the Cooperative Research Centre for Spatial Information through Project 4.09, whose activities are funded by the Australian Commonwealth's Cooperative Research Centres Programme. The Australian Research Council and the Australian Coal Association Research Program have been funding radar related studies by the team at the University of New South Wales (UNSW) during the last few years. The authors wish to thank the Earth Remote Sensing Data Analysis Center (ERSDAC) for providing ALOS PALSAR data. Also thank the European Space Agency (ESA) for providing ASAR data. They acknowledge the strong support from the International Association of Geodesy's Sub-Commission 4.4 "Applications of Satellite & Airborne Imaging Systems" and the IAG Consortium for Mine Subsidence Monitoring. METI and JAXA retain the ownership of the ALOS PALSAR original data. The PALSAR Level-1.1 products were produced and provided to the CRC-SI/UNSW by ERSDAC, Japan.

References

- China Earthquake Administration (2008) [Online] <http://www.cea.gov.cn>. Accessed 23 December 2008
- CMONOC (2008) GPS observations of the co-seismic deformation field for the 2008 Ms 8.0 Wenchuan Earthquake. *Sci China Ser D Earth Sci* 38:1195–1206
- Demos Y-L, Buck C, Guijarro J et al (2000) The ENVISAT advanced synthetic aperture radar system. In: *Proceedings of IGARSS 2000, Honolulu, USA, Vol. 3*, pp 1171–1172
- epochtimes.com (2008) Experts estimate over \$75 billion economic loss from Sichuan Earthquake [Online] <http://en.epochtimes.com/news/8-5-26/71022.html>. Accessed 10 November 2009
- Farr T, Rosen P, Caro E et al (2007) The shuttle radar topography mission. *Rev Geophys*, doi:10.1029/2005RG000183 45
- Ferretti A, Savio G, Barzaghi R et al (2007) Submillimeter accuracy of InSAR time series: Experimental validation. *IEEE Trans Geosci Remote Sens* 45:1142–1153
- Fialko Y, Simons M, Agnew D (2001) The complete (3-D) surface displacement field in the epicentral area of the 1999 Mw 7.1 Hector Mine Earthquake, California, from space geodetic observations. *Geophys Res Lett* 28:3063–3066
- Fornaro G, Guarnieri AM (2002) Minimum mean square error space-varying filtering of interferometric SAR data. *IEEE Trans Geosci Remote Sens* 40:11–21
- Gabriel AK, Goldstein RM, Zebker HA (1989) Mapping small elevation changes over large areas: differential radar interferometry. *J Geophys Res* 94:9183–9191
- Ge L, Ng AH-M, Wang H et al (2009) Crustal deformation in Australia measured by satellite radar interferometry using ALOS/PALSAR imagery. *J Appl Geodes* 3:47–53
- Goldstein RM, Werner CL (1998) Radar interferogram filtering for geophysical applications. *Geophys Res Lett* 25:4035–4038
- Goldstein RM, Engelhardt H, Kamb B et al (1993) Satellite radar interferometry for monitoring ice sheet motion: Application to an Antarctic ice stream. *Science* 262:1525–1530

- Jónsson S, Zebker H, Segall P (2002) Fault slip distribution of the 1999 Mw7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements. *Bull Seismol Soc Am* 92:1377–1389
- Lanari R, Lundgren P, Sansosti E (1998) Dynamic deformation of Etna volcano observed by satellite radar interferometry. *Geophys Res Lett* 25:1541–1544
- Lee J-S, Papathanassiou KP, Ainsworth TL et al (1998) A new technique for noise filtering of SAR interferometric phase images. *IEEE Trans Geosci Remote Sens* 36:1456–1465
- Massonnet D, Feigl KL (1995) Satellite radar interferometric map of the coseismic deformation field of the M=6.1 Eureka Valley, California, earthquake of May 17, 1993. *Geophys Res Lett* 22:1541–1544
- Massonnet D, Rossi M, Carmona C et al (1993) The displacement field of the Landers earthquake mapped by radar interferometry. *Nature* 364:138–142
- Ng AH-M, Chang HC, Ge L et al (2009) Assessment of radar interferometry performance for ground subsidence monitoring due to underground mining. *Earths Planets Space* 61:733–745
- Rosenqvist A, Shimada M, Watanabe M (2007) ALOS PALSAR: A pathfinder mission for global-scale monitoring of the environment. *IEEE Trans. Geosci Remote Sens* 45:3307–3316
- sina.com.cn (2008) Casualties of the Wenchuan Earthquake [Online] <http://news.sina.com.cn/pc/2008-05-13/326/651.html>. Accessed 10 November 2009
- Takeuchi S, Yamada S (2002) Comparison of InSAR capability for land subsidence detection between C-band and L-band SAR. In: *Proceedings of IGARSS 2002, Toronto, Canada, Vol. 4*, pp 2379–2381
- Wright T, Parsons B, Lu Z (2004) Toward mapping surface deformation in three dimensions using InSAR. *Geophys Res Lett* 31:1
- Yun S, Segall P, Zebker HA (2006) Constraints on magma chamber geometry at Sierra Negra volcano, Galapagos Islands, based on InSAR observations. *J Volcanol Geotherm Res* 150: 232–243
- Zebker HA, Rosen PA, Goldstein RM et al (1994) On the derivation of coseismic displacement-fields using differential radar interferometry – the Landers Earthquake. *J Geophys Res* 99:19617–19634

Chapter 18

Uncovering the Space–Time Patterns of Change with the Use of Change Analyst – Case Study of Hong Kong

Bo Huang and Hua-Leung Sin

Abstract Land use change is a complex process and a variety of factors are involved. Recently, a new field, i.e., post-change detection analysis or change analysis, is emerging as a further step after change detection which has been extensively studied. The study of this paper aims to uncover the spatio-temporal pattern of urban development related to the extension of transportation networks in Hong Kong from 1991 to 2007 including the rail and road system as an application of Change Analyst, a spatial statistics tool designed and developed for land-use change modeling. Change Analyst is applied in this paper to perform logistic regression to analyze change patterns and predict their future trend. The results show that the urbanization process in Hong Kong, particularly the non-built-up to built-up land conversion, is highly associated with the transportation network, surrounding land use, and proximity to the central business district that basically refers to the accessibility from the city center. It is also demonstrated that Change Analyst is a powerful tool which assists in various land use change analyses and predictions.

18.1 Introduction

Urbanization is a term to describe urban growth. According to United Nations Population Division of 2005 there were about 3.2 billion people living in urban area globally, which accounted for 48.7% of total population of the world (UNPD 2005). Further, the United Nations (UN) has projected that half of the world population would live in urban areas at the end of 2008 (UN News). With the increasing number of people living in urban areas, the urbanization process has become one of the prominent global issues.

There are two inter-related interpretations of urbanization – (1) physical growth of urban area and (2) movement of people from rural to urban areas. In this study, we

B. Huang (✉)

Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China
e-mail: bohuang@cuhk.edu.hk

focus on the physical growth of urban area. Physical growth of urban area basically refers to the expansion of cities. The major driven cause of physical urban growth is the population increase in the urban area, which includes both internal population growth and rural urban migration. As a result of population increase, more space is needed and people would move to the periphery of a city, turning the non-urban land into urban land, which leads to the physical growth of urban area. In most cases this growth occurs in rural area and this process is therefore known as rural/urban conversion.

Rural/urban conversion has been taken place at the periphery of cities around the globe. Although the growth of urban area (and urban population) provides room for economic development of a city, numbers of impacts do come along with the conversion process such as impact on agricultural production, heat island phenomena, increasing demand of food, water, and sewerage, etc. All these impacts have to be solved by the decision makers. In order to plan ahead of time, understanding urbanization and modeling urban growth become valuable for the decision making process. Therefore, there is a need to strike a balance between the urban growth process and these impacts.

In recent decades, change detection has been studied intensively in different areas of remote sensing and a number of algorithms have been developed to monitor the land-use/land-cover change (Mas 1999; Coppin et al. 2003; Wu 1998). Change detection has already answered the question of “where” and “what” have been changed. A new field of research, post-change detection analysis, is therefore emerging as a further step to answer the question of “how” and “why” the change occurs.

Land use change is a complex issue which is influenced by various factors (such as geographic, political, economic, ecological, etc.). There is no universal rule of how these factors influence land use change and each urbanization process around the world is unique in terms of the combination and weighting of a set of factors.

Developed by the first author’s group, Change Analyst, an extension to ArcGIS, is a GIS tool for land use change modeling and analysis. Using Change Analyst, researchers are able to build up a model of land-use change in relation to various causal factors and to predict the future land-use patterns.

As a case study of Change Analyst, this paper aims to uncover the special-temporal pattern of urban development primarily related to the extension of transportation networks in Hong Kong from 1991 to 2007.

In this chapter, Section 18.2 presents a review of related land use change models. An introduction of study area and data preparation is presented in Section 18.3. Sections 18.4 and 18.5 describe the methodology and the results and discussions, respectively. Finally, the paper is summed up with a conclusion in Section 18.6.

18.2 Land Use Change Modeling

Land use change is a complex process. The non-urban land use to urban land-use conversion, i.e., urban growth, is influenced by a number of socio-economic, political, and environmental factors. Many previous studies have focused on the

effects of various driving forces, including natural environmental characteristics, demographic factors, economy, transportation, preference for proximity, neighborhoods, and government policies (e.g. Cervero and Wu 1997; Mayer and Somerville 2000; Smersh et al. 2003; Angel et al. 2005).

Some of the previous studies incorporated these variables to construct bi-temporal models. For instance, Landis and Zhang (2000) constructed a model with four categories of variables incorporated, which include transportation network, urban structure, locations available for change, and locations where changes were impossible. Verburg et al. (2001), on the other hand, incorporated demographic data in the study. These studies all tried to construct a model of urban growth with multiple categories of variables incorporated and to project future urbanization scenarios.

With regard to model construction, there are a number of techniques that have been studied by researchers. These techniques include multiple regression (Theobald and Hobbs 1998), Markov chain analysis (Lopez et al. 2001; Weng 2002), cellular automata (CA) (Clarke and Gaydos 1998; Wu 1998; Batty et al. 1999; Li and Yeh 2001), and logistic regression (Wu and Yeh 1997; Cheng and Masser 2003). However, there are still limitations of these techniques. For instance, multiple regression cannot ensure high generalization performances for future land-use change projection; Markov chain analysis lacks explanatory power; Cellular automata is a time-consuming process and it is not focused on the interpretation of land-use change patterns (Huang et al. 2009).

Among these techniques, logistic regression is considered a more effective tool for land-use change analysis and interpretation (Cheng and Masser 2003; Munroe et al. 2004; Páez and Suzuki 2001). Logistic regression can produce a functional relationship between the land use change and the driven forces represented by a set of selected explanatory variables. The significance of these variables generated can be used for further interpretation of the weightings of different variables. However, Huang et al. (2009) also reported that multi-temporal changes have not been considered in many of the post-change detection analysis studies, which became a crucial problem to be solved in land-use change modeling.

In order to circumvent this problem, we have developed a multi-temporal spatial regression model for land-use change analysis. In this paper, this regression model under the support of Change Analyst, an extension to ArcGIS, is applied to study the urban growth in Hong Kong.

18.3 Study Area

Hong Kong Special Administrative Region, commonly known as Hong Kong, is located at the eastern part of the Pearl River Delta mouth (Fig. 18.1). The total area is around 1104 km² and the population is around 6.7 million. Hong Kong is one of the most developed cities of China and one of the biggest cities of the world.



Fig. 18.1 Location of Hong Kong

Source: http://en.wikipedia.org/wiki/File:Hong_Kong_Location.svg

18.4 Methodology

In order to interpret the spatial-temporal patterns of land-use change, Change Analyst is used. Change Analyst provides both binary change modeling and multinomial change modeling capabilities. This paper adopts binary change modeling for analysis.

18.4.1 Logistic Regression

Logistic regression has been widely applied in model construction with a categorical dependent variable and both continuous and categorical independent variables. In this paper, a binary logistic regression model is used to analyze the patterns of land-use change with the driving force of transportation network system and to generate an urban growth prediction map. The probability of the land-use change from non-built-up area (value of 0) to built-up area (value of 1) is assumed to follow the logistic curve and the general form of logistic regression is as follows:

$$y = a + b_1x_1 + b_2x_2 + \cdots + b_mx_m \quad (18.1)$$

$$y = \log_e \left(\frac{p}{1-p} \right) = \log \text{it}(P) \quad (18.2)$$

$$P(z = 1) = \frac{e^y}{1 + e^y} \quad (18.3)$$

where x_1, x_2, \dots, x_m are explanatory variables and the utility function y represents a linear relationship between x and y . The parameters b_1, b_2, \dots, b_m refer to the regression coefficients. z is a binary response variable (0 or 1) (0 refers to no change and 1 refers to a new unit, such as the transition from a rural unit to an urban unit).

P refers to the probability of occurrence of a new unit, i.e., $z = 1$. Function y is represented as $\text{logit}(P)$, i.e., the log (to base e) of the odds or likelihood ratio that the dependent variable z is 1. P increases when the y value increases.

Due to the fact that change usually occurs in clusters, spatial dependence should be considered to avoid unreliable parameter estimation. There are two approaches of implementation:

To build a model with autoregressive structure incorporated (Anselin 1988). For example, a spatial lag model can be used as an extension of Eq (18.1):

$$y = a + \sum_{i=1}^m x_i b_i + \rho W_y + \varepsilon \quad (18.4)$$

where ρ is a coefficient on the spatially lagged dependent variable and W is a spatial weight matrix. The maximum likelihood estimator (MLE) is usually adopted to derive the parameters of such a model that best fit the data:

$$L = y \ln \frac{\exp(a + X\beta + \rho W_y)}{1 + \exp(a + X\beta + \rho W_y)} - (1 - Y) \ln[1 + \exp(a + X\beta + \rho W_y)] \quad (18.5)$$

The likelihood can be maximized using a simplex unvaried optimization routine (LeSage 1998).

Spatial Autologistic Regression model or Spatial AutoLogit (SAL) model is the logistic version of Eq. (18.4). It incorporates spatial autocorrelation (Dubin 1995, 1997; LeSage 1998). Such models are effective in regression with spatial autocorrelation considered (Páez and Suzuki 2001). Similar to the spatial lag model, a spatial error model (Anselin 1988) can also be utilized to deal with spatial autocorrelation in a regression model. Nonetheless, they are primarily used for diagnostic analysis rather than for extrapolation-like prediction (Jetz et al. 2005).

To filter out spatial autocorrelation (Getis and Griffith 2002). This can be done by a spatial sampling scheme for spatial autocorrelation reduction among the sampled sites (Munroe et al. 2004). However, a smaller sample size might lead to some information loss and conflicts with the large samples of asymptotic normality using the maximum likelihood method. This issue, yet, can be minimized by appropriate design of the spatial sampling scheme.

In this study the approach b is adopted and the sampling scheme will be discussed in the next section of this paper.

18.4.2 Spatial Sampling

Spatial sampling aims to filter out spatial autocorrelation. Theoretically, spatial autocorrelation should be subject to distance decay. The spatial sampling in this study uses a non-overlapping moving window to verify the spatial autocorrelation of observations proximate in space. A check of joins is determined each time by comparing land use change types of “adjacent” cells. A join is sequential occurrences of like land use changes in adjacent cells (the central cells of the current

sampling window and the sampling windows to the east, west, north, and south of this window).

The size of the sampling window is critical. A small sampling window is insufficient for removing spatial autocorrelation and a large sampling window will result in the loss of information and conflicts with the large-sample of asymptotic normality of the maximum likelihood method for its smaller sample size. In this study, a sampling window size of 5x5 is selected after comparing different sample window sizes.

18.4.3 GIS-Based Predictor Variables

There are six predictor variables considered in the land-use change model (Table 18.1). These variables include sequential land use (built-up area) data and transportation data. There are several underlying assumptions behind these selected variables. First, it is assumed that whether a place is “urban” or not is highly correlated with the accessibility of that place, which basically assumes a possible relationship between transportation and urbanization. Therefore, transportation related variables (such as road, rail network, and distance of a place to the central business district (CBD) are included. Second, it is assumed that built-up area usually develops in a place where there are already some building-area existed nearby, this is why the variable of building density is included. At last, it is also assumed that a built-up area would develop more likely in a gentle slope area rather than a steep slope area. Therefore, the variable of slope is also included in the study. The spatial distribution of these variables is shown in Fig. 18.2.

Table 18.1 Summary of predictor variables for the land use change model

Variables	Definition
Dis_road	Distance from the cell to the nearest major road
Dis_Rail	Distance from the cell to the nearest rail station
R_Den	Density of major road
B_Den	Density of surrounding built-up area cell
Dis_CBD	Distance from the cell to the central business district
Slope	Measurement of the degree of slope

18.4.4 Data Compilation

The data used in this study include land-use data (built-up area and non-built-up area), terrain data, and transportation network data over the time periods of 1991–1997, 1997–2001, and 2001–2007. The land-use data is extracted from the maps of Hong Kong produced by the Lands Department and the former Building and Lands Department of the Hong Kong government. The scale of these maps is

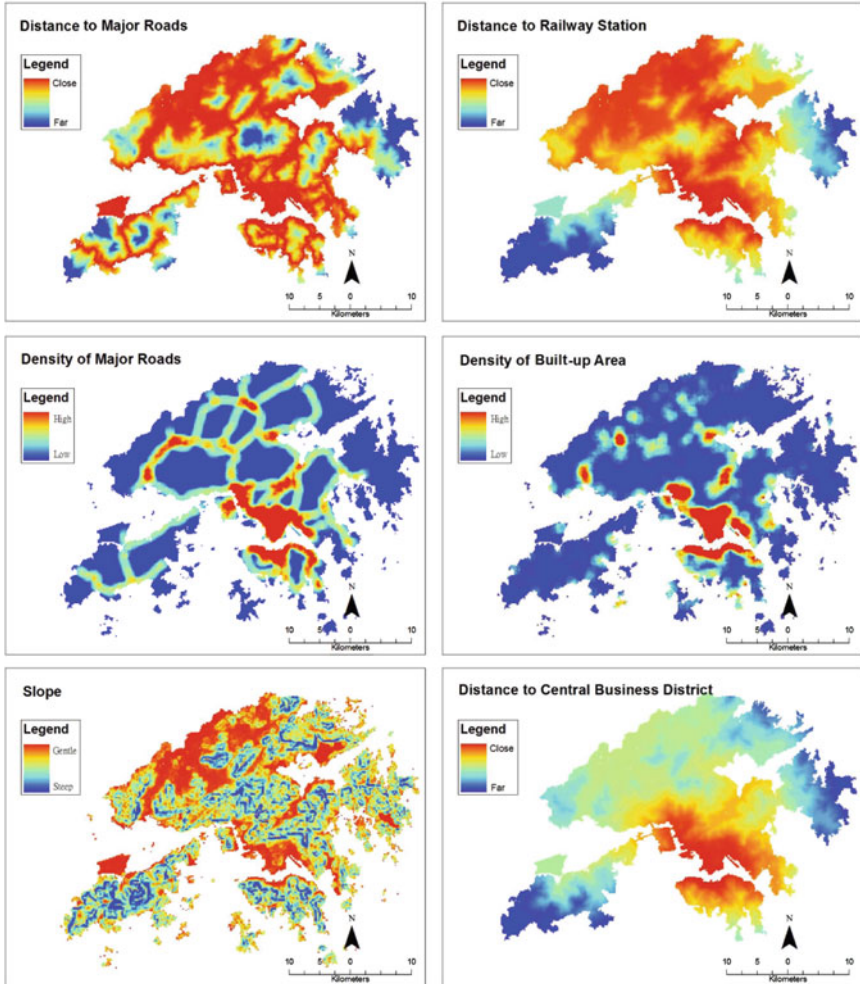


Fig. 18.2 Spatial distribution of predictor variables of the data of 1997

1:200,000 except the 1991 map, which is in 1:100,000. All the data are converted into raster format with the cell size of 180×180 .

The land-use data include built-up area and non-built-up area. A category scan is conducted to detect the land-use change and a bi-temporal change maps is generated (Fig. 18.3). Table 18.2 shows the built-up area and non-built-up area in form of cells in different years. Overall the percentage of built-up area in Hong Kong is small (due to large area of hilly area in Hong Kong). The process of non-built-up and built-up area conversion is slow and stable from 1991 to 2007.

The transportation network data are based on the maps published by the Lands Department and the chronological development of rail system in Hong Kong. In the

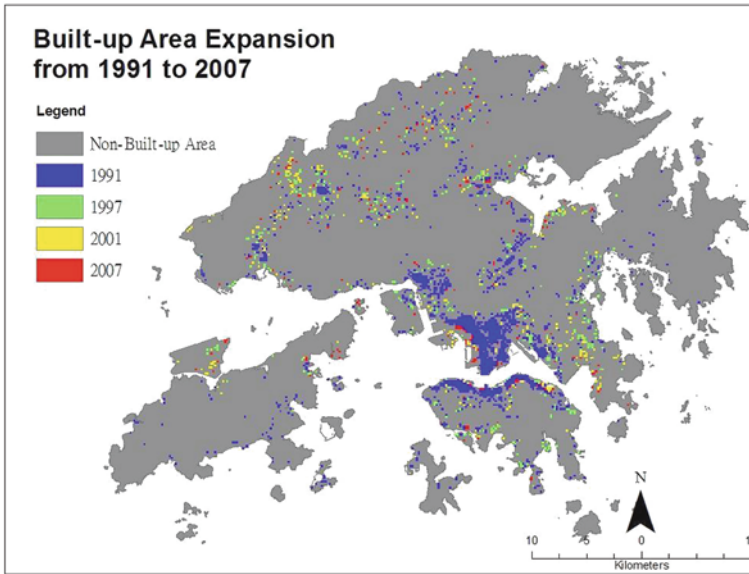


Fig. 18.3 Built-up area and non-built-up area of Hong Kong in different years

Table 18.2 Non-built-up area and built-up area from 1991 to 2007

Categories	1991		1997		2001		2007	
	Cells	%	Cells	%	Cells	%	Cells	%
Non-built-up area	32,513	95.3	32,132	94.2	31,837	93.3	31,660	92.8
Built-up area	1606	4.7	1987	5.8	2281	6.7	2457	7.2

paper maps published by the Lands Department, there are two categories of roads: major road and secondary road. Only the major roads defined by the raw maps are considered in this study.

18.5 Results and Discussion

18.5.1 Logistic Regression Results

The results of the binary logistic regression model are shown in Table 18.3. All the models are statistically significant with the p value close to 0. The percentages of correct prediction (PCP) are 97.96% for the period of 1991–1997, 98.51% for the period of 1997–2001 and 99.41% for the period of 2001–2007.

Table 18.3 Regression results

Variable	Model 1991–1997		Model 1991–2001		Model 1991–2007		Model 1997–2001		Model 1997–2007		Model 2001–2007	
	Coefficient	<i>t</i> -probability	Coefficient	<i>t</i> -probability	Coefficient	<i>t</i> -probability	Coefficient	<i>t</i> -probability	Coefficient	<i>t</i> -probability	Coefficient	<i>t</i> -probability
Dis_Rail	-3.945995	0.0996	-2.131598	0.2972	-1.182973	0.5425	2.176413	0.5796	0.893402	0.7549	-2.154502	0.7418
R_Den	2.791800	0.4107	-2.158267	0.4423	-0.695555	0.7791	-10.191364	0.0115	-6.399769	0.0381	-1.361451	0.7663
Dist_Road	9.906431	0.2057	14.797678	0.0359	2.731206	0.3988	23.345911	0.0868	2.692749	0.5807	-4.577834	0.4222
Slope	7.695564	0.0107	4.690655	0.0269	7.112473	0.0003	0.030267	0.9916	5.112566	0.0458	10.682266	0.0344
Dis_CBD	6.296543	0.0025	5.398179	0.0008	5.315806	0.0003	2.378973	0.3901	3.297601	0.1190	6.215349	0.0681
B_Den	-2.907837	0.3343	0.279119	0.9058	0.957930	0.6471	7.707397	0.0123	6.628498	0.0098	2.989448	0.4496
Constant	-21.4368	—	-23.5239	—	-14.8145	—	-30.4078	—	-14.1353	—	-12.9158	—
<i>Models description</i>												
Threshold Value	0.15		0.15		0.15		0.15		0.15		0.15	
<i>p</i> value	0.000001		0.000000		0.000000		0.000045		0.000001		0.009765	
PCP	97.961%		95.677%		94.78%		98.51%		97.351%		99.414%	

18.5.2 Evaluation of the Model

To testify the capability of the model, a prediction land-use map of 2003 is produced by the model and compared to the actual land-use map of 2003 (Fig. 18.4 for the map and Table 18.4 for the comparison of statistics).

18.5.3 Prediction Results

The following figure (Fig. 18.5) is the predicted land-use map of 2015 Hong Kong with the use of Change Analyst and the statistics of the predict 2015 land-use map are shown in Table 18.5.

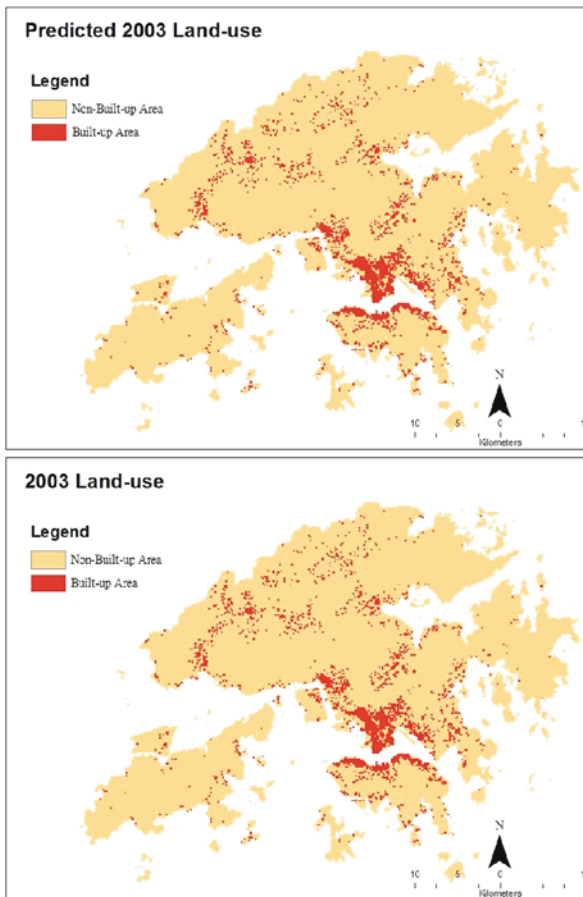


Fig. 18.4 Predicted and actual land-use map of 2003

Table 18.4 Comparison of statistics between predicted and real land-use data of 2003

		Predicted land-use of 2003		
		Non-built-up area (0)	Built-up area (1)	Total
Real land-use of 2003	Non-built-up area (0)	31,739	15	31,754
	Built-up area (1)	67	2297	2364
	Total	31,806	2312	34,118
PCP		99.8%	99.4%	99.8%

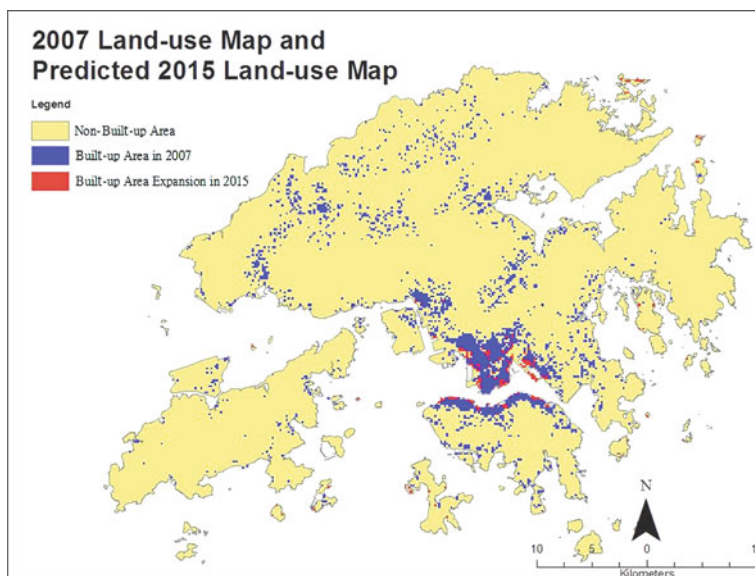


Fig. 18.5 Predict 2015 land-use map

Table 18.5 Statistics of predicted 2015 land-use map

	2015	
	Cell	%
Non-built-up area	31,441	92.2
Built-up area	2676	7.8

According to the predicted 2015 land-use map, most of the expended Built-up Areas in 2015 compared to 2007 are located in Kowloon peninsula, mainly east Kowloon and west Kowloon (including the area of Kowloon Bay, Hung Hum, the runway of previous Kai Tak Airport, Mei Foo, and the new reclamation land of west Kowloon).

18.6 Discussion

The study of this paper selects transportation data as the major component of the explanatory variables and aims to uncover the pattern of urban growth in Hong Kong in relation with transportation networks. There are some trends of non-built-up and built-up area conversion over time in Hong Kong according to the results of this study.

The speed of urban growth in Hong Kong is slow and stable. It is probably the result of a relatively high urbanization rate before the 1990s and most of the developable land has been utilized already. Land is limited in Hong Kong and the urban growth is also limited (spatially) as well.

According to the regression results, the variables of slope, distance to CBD and distance to major roads contribute a significant and relatively stable influence over time in the logistic regression model. The result is reasonable. Because of the natural mountainous topology, slope therefore because a critical factor influencing the mode of urban growth in Hong Kong. The CBD of Hong Kong is Tsim Sha Tsui and Central area. In general, urban area developed around the CBD initially. With all other factors as the same, a place closer to the CBD is assumed to be more likely urban than a place located away from the CBD. The result of the model has proved this assumption to be applicable in Hong Kong. Distance to Major Roads is another significant variable in the model. Basically this variable is related to the accessibility of a place. In Hong Kong, road access usually has to be completed before development to allow public transportation going into the area in the future. Therefore, distance to major roads is also a critical factor as in the results of the regression model.

However, the factor of distance to railway station is unstable in the model. It may be the result of the railway policy and urban growth mode in Hong Kong. The identification of railway station location in Hong Kong is not random. A station cannot operate until the population of the area reaches a certain amount which proves that there is a need for a massive transportation (railway system). However, the area usually has already been developed to be a built-up area before the rail station starts its operation. Instead of the non-built-up to built-up area conversion, the influence of a new railway station will probably stimulate the land price increase and “re-urbanize” the area (demolish the old building and replace it by new buildings). Therefore, the nature of built-up area does not change.

Overall, the findings of the regression models are consistent and reasonable with the pattern of urban expansion in Hong Kong.

18.7 Conclusion

Land use change analysis is a highlighted issue around the world especially in the face of global change. While urbanization is undergoing in many areas around the world, planning for better utilization of space (with the understanding that space is

limited) is one of the major concerns of policy makers, academics, and the general public. Change Analyst serves to be a powerful tool to provide a prediction of land use change for the users. As the interface of Change Analyst is user-friendly, users do not need to have any background of logistic regression techniques. For example, policy makers can have a general scope of the future land use trend with the available information set to be the predicting variables. Therefore, they can have a better idea the change in the future and make relevant and appropriate policies. Education is another example showing its applicability. For instance, in geography class, students can use the software to understand the temporal–spatial pattern of land use change.

This study applies a logistic regression model and aims to uncover the spatio-temporal patterns of land use change with Hong Kong as the case study. In Change Analyst, the logistic regression model analyzes how the selected driven forces influence the conversion from non-built-up to built-up area. In addition, the regression model also produces a predicted land-use map in the future. In general, an interpretation can be drawn that slope, the distance to CBD, and the distance to major roads are the major driven forces of urban growth in Hong Kong. This result is reasonable and matches with the current situation of Hong Kong. The results of the analysis show that the logistic regression model can achieve a good performance in land-use change modeling and prediction.

However, data constraint is a key limitation. It is generally impossible to incorporate all variables for modeling. Simplification and generalization processes have to be undertaken, and so loss of data and certain extent of generalization do occur during the model construction. Therefore, a careful design and collection of data is required for the modeling in order to have a more ideal result and a better interpretation. In addition, the current version of Change Analyst cannot deal with the problem of multicollinearity which may occur when explanatory variables are highly correlated. However, it will be solved in the new version(s) of the software in the future.

References

- Anselin L (1988) *Spatial econometrics: methods and models*. Kluwer Academic, Dordrecht
- Angel S, Sheppard SC, Civco DL (2005) *The dynamics of global urban expansion*. Transport and Urban Development Department, The World Bank, Washington, DC
- Batty M, Xie Y, Sun Z (1999) Modeling urban dynamics through GIS-based cellular automata. *Comput Environ Urban Syst* 23:205–233
- Clarke KC, Gaydos LJ (1998) Loose-coupling a CA model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *Int J Geogr Inf Sci* 12:699–714
- Cervero R, Wu K-L (1997) Polycentrism, commuting, and residential location in the San Francisco Bay Area. *Environ Plan A* 29:865–886
- Cheng J, Masser I (2003) Urban growth modeling: a case study of Wuhan city, PR China. *Landscape Urban Plan* 62:199–217
- Coppin P, Jonckheere I, Nackaerts K, Muts B (2003) Digital change detection methods in ecosystem monitoring: a review. *Int J Remote Sens* 24:1–33
- Dubin R (1995) Estimating logit models with spatial dependence. In: Anselin L, Florax R (eds) *New directions in spatial econometrics*. Springer-Verlag, New York, pp 229–242

- Dubin R (1997) A note on the estimation of spatial logit models. *Geogr Syst* 4:181–193
- Getis A, Griffith DA (2002) Comparative spatial filtering in regression analysis. *Geogr Anal* 34:130–140
- Greene W (2000). *Econometric analysis*. Prentice-Hall, Upper Saddle River
- Huang B, Zhang L, Wu B (2009) Spatiotemporal analysis of rural–urban land conversion. *Int J Geogr Inform Sci* 23:379–398
- Jetz W, Rahbek C, Lichstein JW (2005) Local and global approaches to spatial data analysis in ecology. *Global Ecol Biogeogr* 14:97–98
- Landis JH, Zhang M (2000) Using GIS to improve urban activity and forecasting models: three examples. In Fotheringham AS, Wegener M (eds) *Spatial models and GIS – new potential and new models*. Taylor & Francis, London, pp 63–81
- LeSage JP (1998) *Learning materials on spatial econometrics*. Department of Economics, University of Toledo, Toledo, OH
- Li X, Yeh AG (2001) Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environ Plan A* 33:1445–1462
- Lopez E, Bocco G, Mendoza M, Duhau E (2001) Predicting and cover and land use change in the urban fringe: a case in Morelia city, Mexico. *Landscape Urban Plan* 55:271–285
- Mas JF (1999) Monitoring land-cover changes: a comparison of change detection techniques. *Int J Remote Sens*, 20:139–152
- Mayer CJ, Semerville CT (2000) Land use regulation and new construction. *Reg Sci Urban Econ* 30:639–662
- Munroe DK, Southworth J, Tucker CM (2004) Modeling spatially and temporally complex land-cover change: the case of western Honduras. *Prof Geogr* 56:544–559
- Páez A, Suzuki J (2001) Transportation impacts on land use change: an assessment considering neighborhood effects. *J Eastern Asia Soc Transport Stud* 4:47–59
- Smersh GT, Smith MT, Schwartz AL (2003) Factors affecting residential property development patterns. *J Real Estate Res* 25:61–75
- Theobald DM, Hobbs NT (1998) Forecasting rural land use change: a comparison of regression and spatial transition-based models. *Geogr Environ Model* 2:65–82
- UN News. UN says half the world's population will live in urban areas by end of 2008. <http://www.un.org/apps/news/story.asp?NewsID=25762&Cr=population&Cr1=#>. Accessed on 25 June, 2009
- UNPD (United Nations Population Division) (2005) *World population prospects: the 2004 revision and world urbanization prospects: the 2005 revision*. United Nations, New York
- Verburg PH, Koning, GHJ, Kok K, Veldkamp A., Priess J (2001) The CLUE modelling framework: an integrated model for the analysis of land use change. In Singh RB, Jefferson F, Himiyama Y (eds). *Land use and cover change*. Science Publishers, Enfield, NH
- Weng Q (2002) Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS, and stochastic modeling. *J Environ Manage* 64:273–284
- Wu F (1998) SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *Int J Geogr Inform Sci* 12:63–82
- Wu F, Yeh AG (1997) Changing spatial distribution and determinants of land development in Chinese cities in the transition from a centrally planned economy to a socialist market economy: a case study of Guangzhou. *Urban Stud* 34:1851–1879

Chapter 19

Change Detection of Sea Ice Distribution in SAR Imagery Using Semi-variogram of Intrinsic Regionalization Model

Yu Li and Jonathan Li

Abstract The spatial structures revealed in remotely sensed imagery are essential information characterizing the nature and the scale of spatial variation of sea ice processes. This study evaluates the potential capability of using semi-variogram of intrinsic regionalization model for change detection of sea ice. Up to now, the second-order variogram has been widely used to describe the spatial variations within an image, but it demonstrates the limitation to discriminate distinct image spatial structures. This study introduces a different geo-statistic metric, in which spatial structures of sea ice are considered a combination of two stochastic second-order stationary models. Firstly, the multi-gamma model is used to characterize continuous variations corresponding to water or the background of sea ice. The second model is a tessellation model, in which the image domain is randomly separated into non-overlapping cells. In each cell, a random value is independently assigned. It is called the mosaic model. In this paper, the mosaic model is constructed by a Poisson tessellation. The linear combination of these two stochastic models defines the mixture model to represent spatial structures of sea ice presented in SAR intensity imagery. This algorithm is applied to Radarsat-1 images acquired different days to identify the change of sea ice.

19.1 Introduction

Sea ice imposes severe restriction on ship traffic in the Arctic, is a sensitive climate indicator, and plays an important role in exploration and exploitation of marine resources (Johannessen et al. 2007). To evaluate the impact of sea ice on ship traffic and global climate system, it is necessary to observe and monitor its change daily, seasonally and yearly. However, extreme remoteness and harsh environments

Y. Li (✉)

Department of Geography and Environmental Management, Faculty of Environment,
University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1
e-mail: y62li@uwaterloo.ca

of Polar Regions make fieldwork very expensive and impossible in many cases. Given this challenge, remote sensing has proved to be a powerful tool for monitoring sea ice. In particular, synthetic aperture radar (SAR) systems have been shown to be very useful due to its capability of acquiring data under all weather condition, during day and night, and over large region (Kerman 1999). SAR imagery may exhibits the spatial heterogeneity of the retrieved sea ice property. This information helps to characterize the nature of the processes structuring sea ice, identifying its scale of spatial variation. This study focuses specifically on the characterization of the spatial structures of sea ice presented in SAR data. These spatial structures are specific to the measured change of sea ice.

In the next section the intrinsic model is outlined. In Section 19.3 the parameter estimation for the mixture model is described. The results are discussed in Section 19.4 and finally, in Section 19.5 the conclusions are given.

19.2 Intrinsic Model Based on Bigamma and Mosaic Random Functions

19.2.1 Intrinsic Model

From geostatistics point of view, the spatial structure and variability revealed in SAR data can be modeled by a regionalized variable $z(\mathbf{x})$, which is simply considered as the realization of a random function $Z(\mathbf{x})$ constructed at all points \mathbf{x} of a given region $D \subset \mathbb{R}^2$. When the random function is stationary, its moments are obviously invariant under translations, e.g., its first two moments for points at \mathbf{x} and $\mathbf{x} + \mathbf{h}$ can be written as

$$\begin{cases} E[Z(\mathbf{x})] = m \\ E[Z(\mathbf{x}) - m][Z(\mathbf{x} + \mathbf{h}) - m] = C(\mathbf{h}) \end{cases} \quad (19.1)$$

where E is the operator of the mean value, and m is the mean and $C(\mathbf{h})$ is the covariance function. When m is constant $C(\mathbf{h})$ only depends on the vector \mathbf{h} , that is, the random function is second-order stationary (Chiles and Delfiner 1999; Wackernagel 2003). A second-order random function is isotropic if its covariance function only depends on the length $\|\mathbf{h}\|$, denoted by h for convenience, of the vector \mathbf{h} and not depends on its direction, that is, $C(\mathbf{h}) = C(h)$.

Under intrinsic hypothesis, for every vector \mathbf{h} the increment $Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})$ is a second-order random function of \mathbf{x} . Then the intrinsic random function can be characterized by the moments of the increment

$$\begin{cases} E[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})] = 0 \\ E[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2 = 2\gamma(h) \end{cases} \quad (19.2)$$

where $\gamma(h)$ is its semi-variogram function.

Following Wackernagel (2003), the intrinsic random function $Z(\mathbf{x})$ can be divided into a sum of spatial components characterizing different spatial scale. For a two component model

$$Z(\mathbf{x}) = Z_1(\mathbf{x}) + Z_2(\mathbf{x}) \quad (19.3)$$

where the increments for components are zero mean and uncorrelated, that is,

$$\begin{cases} E[Z_i(\mathbf{x} + \mathbf{h}) - Z_i(\mathbf{x})] = 0, \text{ for } i = 1, 2 \\ E[(Z_1(\mathbf{x} + \mathbf{h}) - Z_1(\mathbf{x})) (Z_2(\mathbf{x} + \mathbf{h}) - Z_2(\mathbf{x}))] = 0 \end{cases} \quad (19.4)$$

Based on this idea, the intrinsic random function $Z(\mathbf{x})$ is defined as a weighted sum of two components: bigamma and Poisson mosaic random functions (Garrigues et al. 2007), respectively, to characterize the spatial structures of sea ices in different seasons, that is,

$$Z(\mathbf{x}) = wZ_1(\mathbf{x}) + \sqrt{1 - w^2}Z_2(\mathbf{x}) \quad (19.5)$$

The semi-variogram of the random function can be derived by

$$\begin{aligned} \gamma(h) &= E[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2 \\ &= E\left[w(Z_1(\mathbf{x} + \mathbf{h}) - Z_1(\mathbf{x})) + \sqrt{1 - w^2}(Z_2(\mathbf{x} + \mathbf{h}) - Z_2(\mathbf{x}))\right]^2 \\ &= w^2E[Z_1(\mathbf{x} + \mathbf{h}) - Z_1(\mathbf{x})]^2 + (1 - w^2)E[Z_2(\mathbf{x} + \mathbf{h}) - Z_2(\mathbf{x})]^2 \\ &= w^2\gamma_1(h) + (1 - w^2)\gamma_2(h) \end{aligned} \quad (19.6)$$

where $\gamma_1(h)$ and $\gamma_2(h)$ are the semi-variogram functions of components $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$, respectively and will be defined in the following sections.

19.2.2 Bigamma Random Function

A random function $Z_1(\mathbf{x})$ is said to be a bigamma random function if the vectors $\mathbf{Z} = (Z_1(\mathbf{x} + \mathbf{h}), Z_2(\mathbf{x}))$ for $\mathbf{x}, \mathbf{x} + \mathbf{h} \in D$ are distributed according to a bivariate gamma distribution. The bivariate gamma distribution on \mathbb{R}^2 has been defined in several forms. Most of them exploit various properties of the univariate gamma distribution to construct bivariate families (Kotz et al. 2000).

In this paper, bivariate gamma distribution of random vector $\mathbf{Z} = (Z_1(\mathbf{x} + \mathbf{h}), Z_1(\mathbf{x}))$ in \mathbb{R}^2 is defined by its moment generating function or Laplace transform, which is characterized by an affine polynomial (Barndorff-Nielsen 1980; Barndorff 2006). Given an affine polynomial

$$P(\boldsymbol{\theta}) = 1 + \beta\theta_1 + \beta\theta_2 + \beta^2\rho\theta_1\theta_2 \quad (19.7)$$

where the parameters satisfy the conditions: $\beta > 0$ and $1 > \rho > 0$. Then the moment generating function of \mathbf{Z} can be defined as

$$L(\boldsymbol{\theta}) = E \left(e^{\theta_1 Z_1(\mathbf{x}+\mathbf{h})+\theta_2 Z_1(\mathbf{x})} \right) = (P(-\boldsymbol{\theta}))^{-\alpha} = (1 - \beta\theta_1 - \beta\theta_2 + \beta^2\rho\theta_1\theta_2)^{-\alpha} \tag{19.8}$$

From the defined generating function, it is obvious that $Z_1(\mathbf{x})$ are distribution according to a univariate gamma distribution with shape parameter α and scale parameter β , that is, $Z_1(\mathbf{x}) \sim G(\alpha, \beta)$. The probability density function can be expressed as follows

$$f(Z_1(\mathbf{x}); \alpha, \beta) = \frac{Z_1(\mathbf{x})^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp\left(-\frac{Z_1(\mathbf{x})}{\beta}\right) \tag{19.9}$$

where $\Gamma(\cdot)$ is the gamma function.

The moments of the bigamma distribution can be obtained by differentiating Eq. (19.8). For example, the means and variances can be obtained as

$$E[Z_1(\mathbf{x})] = \left. \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_1} \right|_{\boldsymbol{\theta}=0} = \alpha\beta \tag{19.10}$$

$$\text{Var}[Z_1(\mathbf{x})] = \left. \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_2} \right|_{\boldsymbol{\theta}=0} = \alpha\beta^2 \tag{19.11}$$

The covariance $C(Z_1(\mathbf{x}+\mathbf{h}), Z_1(\mathbf{x}))$ and correlation coefficient $r(Z_1(\mathbf{x}+\mathbf{h}), Z_1(\mathbf{x}))$ of the bigamma random function can also be calculated by

$$\begin{aligned} \text{Cov}[Z_1(\mathbf{x}+\mathbf{h}), Z_1(\mathbf{x})] &= E[Z_1(\mathbf{x}+\mathbf{h})Z_1(\mathbf{x})] - E[Z_1(\mathbf{x}+\mathbf{h})]E[Z_1(\mathbf{x})] = \alpha\beta^2(1-\rho) \end{aligned} \tag{19.12}$$

$$r(Z_1(\mathbf{x}+\mathbf{h}), Z_1(\mathbf{x})) = \frac{\text{Cov}(Z_1(\mathbf{x}+\mathbf{h}), Z_1(\mathbf{x}))}{\sqrt{\text{Var}(Z_1(\mathbf{x}+\mathbf{h}))}\sqrt{\text{Var}(Z_1(\mathbf{x}))}} = 1 - \rho \tag{19.13}$$

In order to characterize the spatial structures, the ρ is assumed as exponential function with h

$$\rho = \rho(h) = 1 - \exp\left(-\frac{3h}{r_1}\right) \tag{19.14}$$

Under this assumption, the $Z_1(\mathbf{x})$ defines a second-order stationary isotropic bigamma random function, its covariance function can be rewritten as

$$\text{Cov}[Z_1(\mathbf{x}+\mathbf{h}), Z_1(\mathbf{x})] = \alpha\beta^2 \exp\left(-\frac{3h}{r_1}\right) \tag{19.15}$$

The second-order semi-variogram $\gamma_1(h)$ is thus an exponential variogram with range r_1 and sill $\alpha\beta^2$ as follows

$$\gamma_1(h) = \alpha\beta^2 \left(1 - \exp\left(-\frac{3h}{r_1}\right) \right) \tag{19.16}$$

19.2.3 Poisson Tessellation Based Mosaic Random Function

According to Rivoirard (1994) and Lantuejoul (2002), a mosaic random field on a domain D can be defined by partitioning the domain D into a tessellation and assigning each cell of the tessellation a value drawn from a distribution such that the values do not depend on the cells and different cells have independent values. For a flexible and convenient tessellation, the Poisson line tessellation (Chiles and Delfiner 1999; Lantuejoul 2002) is used to divide the domain $D \subset \mathbb{R}^2$ into small cells.

A line in \mathbb{R}^2 is specified by two parameters $(a, d) \in [0, 2\pi) \times [0, \infty)$ where a is the direction of the unit vector orthogonal to the line and d is the distance from the line to the origin (see Fig. 19.1a). Poisson line network can be completely defined by the intensity λ of the Poisson point process in the parameter space $[0, 2\pi) \times [0, \infty)$. The Poisson lines in the network hitting the domain D form a partition of D into convex polygons (see Fig. 19.1b).

The gamma distribution defined in Eq. (19.9) is used for the distribution from which the values are drawn for each cell of Poisson tessellation. This defines a random function $Z_2(\mathbf{x})$ following the gamma distribution with the shape parameter α and scale parameter β . Note that the statistical properties of $Z_2(\mathbf{x})$ can be characterized by its spatial distribution.

Given two points \mathbf{x} and $\mathbf{x} + \mathbf{h}$ in D , $p(\mathbf{h})$ denotes the probability that both \mathbf{x} and $\mathbf{x} + \mathbf{h}$ belong to the same cell. In isotropic case, $p(\mathbf{h}) = \rho(h)$, then the second order moment of $Z_2(\mathbf{x})$ can be expressed as (Lantuejoul 2002)

$$E[Z_2(\mathbf{x} + \mathbf{h})Z_2(\mathbf{x})] = \alpha\beta^2 + \alpha^2\beta^2\rho(h) + \alpha^2\beta^2(1 - \rho(h)) \tag{19.17}$$

so that,

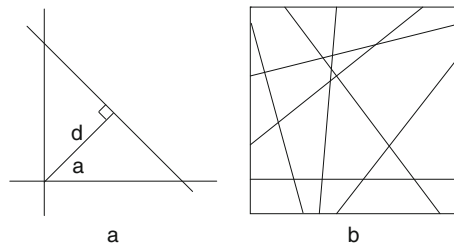


Fig. 19.1 (a) 2D line with parameter (a, d) ; (b) Simulation of Poisson lines and Poisson polygon tessellation

$$\text{Cov}[Z_2(\mathbf{x} + \mathbf{h}), Z_2(\mathbf{x})] = C_2(h) = \alpha\beta^2 p(h) \quad (19.18)$$

For the Poisson mosaic random function, simple probabilistic calculations show (Chiles and Delfiner 1999) that its covariance function is

$$C_2(h) = \alpha\beta^2 \exp(-2\lambda h) \quad (19.19)$$

Similarly, the semi-variogram $\gamma_2(h)$ of $Z_2(\mathbf{x})$ with range r_2 and sill $\alpha\beta^2$ can be obtained

$$\gamma_2(h) = \alpha\beta^2 \left(1 - \exp\left(-\frac{3h}{r_2}\right)\right) \quad (19.20)$$

where $r_2 = \frac{3}{2}\lambda$.

19.3 Parameter Estimation for the Mixture Model

In this section, the least squares estimation is used for the theoretical variogram to estimate the parameters in the model. Hence, an experimental variogram estimator acts as the data for this process.

19.3.1 Experimental Semi-variogram

In order to derive variogram from a set of observations $\mathbf{y} = \{y(x_i); x_i \in D, i = \{1, \dots, n\}\}$, the classical experimental variogram is proposed (Wackernagel 2003; Schabenberger and Gotway 2005)

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} [y(x+h) - y(x)]^2 \quad (19.21)$$

where $N(h) = \{(x_i, x_j); x_i, x_j \in D \text{ and } x_i - x_j = h\}$ and $|N(h)| = \#N(h)$.

The variogram measures the average of squared differences between values $y(x_i)$ and $y(x_j)$ of all pairs of positions (x_i, x_j) separated by a distances $x_i - x_j = h$. The reasons for choosing Eq. (19.21) for the data to estimate the parameters of theoretical variogram, mainly because of its appealing properties: simple computation, unbiased estimation, evenness and attaining zero at zero h (Schabenberger and Gotway 2005). In practice, the experimental variogram is usually computed using vector h with a tolerance ε , that is, pairs of positions $x_i - x_j = h$ that satisfy $x_i - x_j = h \pm \varepsilon$.

19.3.2 Theory Semi-variogram

Substitute Eqs. (19.13) and (19.17) into (19.19)

$$\gamma(|h|) = \alpha\beta^2 \left[1 - \omega^2 \exp\left(-\frac{3|h|}{r_{bg}}\right) - (1 - \omega^2) \exp\left(-\frac{3|h|}{r_{pm}}\right) \right] \quad (19.22)$$

As demonstrated in Eq. (21), the variogram of the proposed mixture model Eq. (18) is determined by four parameters $\theta = (\beta, \omega, r_{bg}, r_{pm})$ supposing that α is a constant equal to the number of looks of SAR sensors. In this paper, those parameters are used to characterize the spatial structure in a SAR intensity image.

19.3.3 Parameter Estimation by Least-Squares Adjustment

Least-squares adjustment technique is usually used to automatically fit a theoretical variogram to experimental one (Cressie 1993; Wackernagel 2003; Schabenberger and Gotway 2005). Given a theoretical variogram model within a family $\gamma(h, \theta)$, where parameter vector θ contains all unknown parameters to be estimated from the data, let $\hat{\boldsymbol{\gamma}}(h) = [\hat{\gamma}(h_1), \dots, \hat{\gamma}(h_k)]^T$ be the values calculated from the $N(h_j)$ available pairs by the experimental variogram estimator Eq. (19.22), consider fitting $\gamma(h, \theta)$ to $\hat{\boldsymbol{\gamma}}(h)$. Suppose that $\boldsymbol{\gamma}(h, \theta) = [\gamma(h_1, \theta), \dots, \gamma(h_k, \theta)]^T$ and $\hat{\boldsymbol{\gamma}}(h)$ satisfy a statistic model of form

$$\hat{\boldsymbol{\gamma}}(h) = \boldsymbol{\gamma}(h, \theta) + \boldsymbol{\mu}(h) \quad (19.23)$$

where $\boldsymbol{\mu}(h)$ is the $k \times 1$ error vector with zero mean and variance-covariance matrix $V(\theta) = \text{Var}[\boldsymbol{\mu}(h)]$.

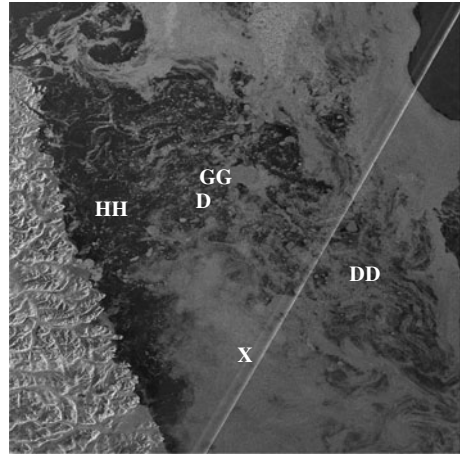
By weighted least squares (Cressie 1993; Schabenberger and Gotway 2005), the parameter vector θ can be estimated so as to minimize the weighted sum of squares

$$\begin{aligned} & (\hat{\boldsymbol{\gamma}}(h) - \boldsymbol{\gamma}(h, \theta))^T W(\theta)^{-1} (\hat{\boldsymbol{\gamma}}(h) - \boldsymbol{\gamma}(h, \theta)) \\ &= \sum_{j=1}^k \frac{|N(h_j)|}{2\gamma(h_j, \theta)} \{ \hat{\gamma}(h_j) - \gamma(h_j, \theta) \} \end{aligned} \quad (19.24)$$

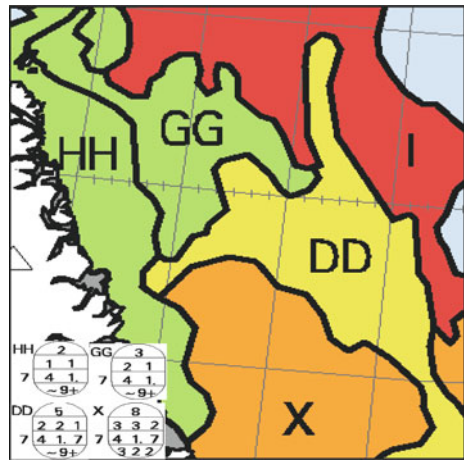
19.4 Experimental Results

The proposed algorithm is tested by using SAR intensity images to identify sea ice structures. Figure 19.2a shows the 4-looks Radarsat-1 intensity images with size of 2159 2941 pixels, vertical-vertical (VV) polarization and 30 m spatial resolution, which is acquired on 12 May 2008 from Labrador Sea, Canada. The area shown

Fig. 19.2 (a) The study area and (b) regional ice analysis and egg codes for the study area

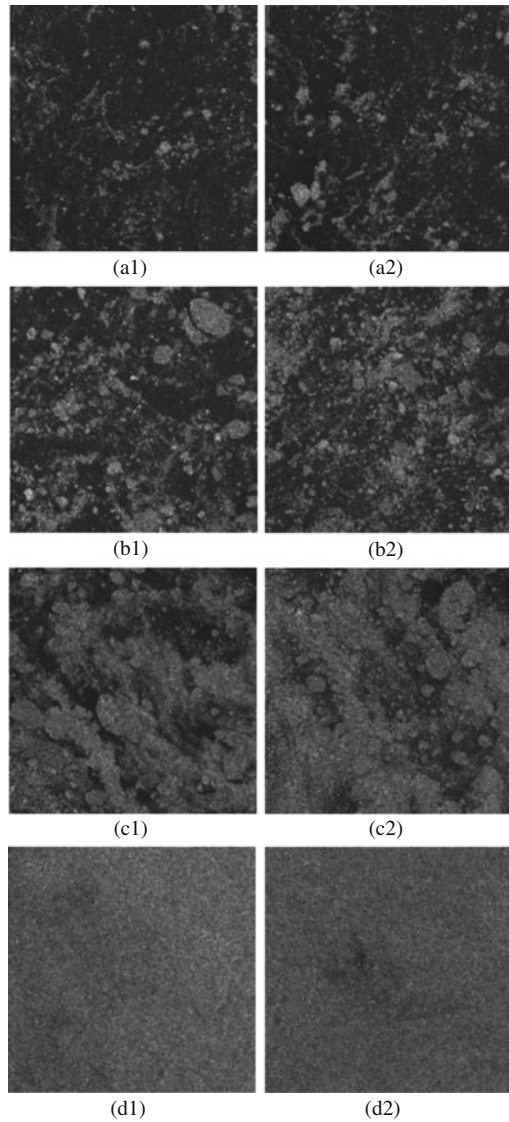


(a)



(b)

in Fig. 19.2 includes four sea ice structures (DD, GG, HH, and X). Figure 19.2b shows the regional ice analysis and the egg codes for these structures provided by Canada Ice Service (CIS), Environment Canada. For each structure, two images are cut for the testing images (see Fig. 19.3), where a–d indicate DD, GG, HH, and X regions, respectively. Figure 19.4 gives the theoretical and experimental semi-variograms for testing images shown in Fig. 19.3. Both theoretical and experimental semi-variograms are normalized. Table 19.1 lists the estimated parameters where w is the weight of two components in intrinsic regionalization models, which can be

Fig. 19.3 Testing images

to indicate the density of sea ice, d is the density of sea ice in the same regions from the egg code, r_1 indicates the global texture structure of water, r_2 is equal to $3\lambda/2$ and λ can be used to estimate the average size of sea ice sheet and the means of estimated parameter are calculated. The results promise w as density indicator.

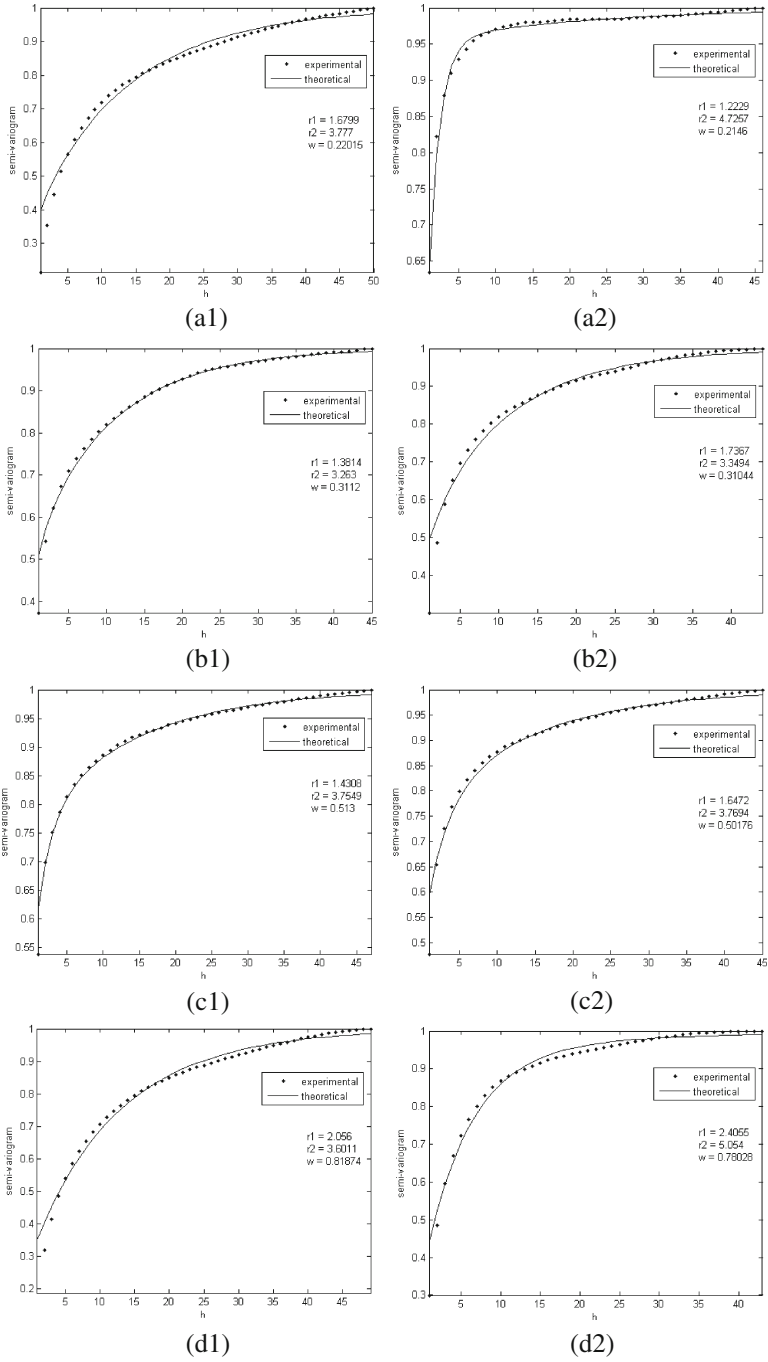


Fig. 19.4 Theoretical and experimental semi-variograms

Table 19.1 Estimated parameters

Structure	Image	r_1	r_2	w	\bar{r}_1	\bar{r}_2	\bar{w}	d
HH	a1	1.7	3.8	0.22	1.5	4.2	0.21	0.2
	a2	1.2	4.7	0.21				
GG	b1	1.4	3.3	0.31	1.5	3.3	0.31	0.3
	b2	1.7	3.3	0.31				
DD	c1	1.4	3.8	0.51	1.5	3.8	0.50	0.5
	c2	1.7	3.8	0.50				
X	d1	2.1	3.6	0.82	2.3	4.3	0.80	0.8
	d2	2.4	5.0	0.78				

19.5 Conclusions

In this study, spatial structures of sea ice revealed in a SAR intensity image are characterized by intrinsic regionalization model. The algorithm is applied on Radarsat-1 SAR intensity image of Ungava Bay, Canada in order to detect the sea ice change during twenty days in May 2008. The results demonstrate that the algorithm is useful to detect sea ice change in intensity and size. However, this result is limited by the number, the types and the small size of the sea ice. Several points of the proposed algorithm can be improved by further studies, such as using the Voronoi mosaic model instead of the Poisson line model.

Acknowledgments The Radarsat-1 sea ice SAR images were provided by Canadian Ice Service (CIS), Environment Canada. The authors would like to thank the anonymous reviewers and Dr. Emilio Chuvieco for their time and effort in reviewing the earlier version of the manuscript, which helped improve the scholarly quality of this paper.

References

Barndorff-Nielsen OE (1980) Conditional resolutions. *Biometrika* 67:293–310

Bernardoff P (2006) Which multivariate gamma distributions are infinitely divisible? *Bernoulli* 12(1):169–189

Chiles J, Delfiner P (1999) *Geostatistics modeling spatial uncertainty*. Wiley, New York

Cressie NAC (1993) *Statistics for Spatial Data Revised Edition*. John Wiley & Sons, New York

Garrigues S, Allard D, Baret F (2007) Using first- and second-order variograms for characterizing landscape spatial structures from remote sensing imagery. *IEEE Trans Geosci Remote Sens* 45(6):1823–1834

Johannessen OM, Alexandrov VY, Frolov IY, Sandven S, Pettersson LH, Bobylev LP, Kloster K, Smirnov VG, Mironov YU, Babich NG (2007) *Remote sensing of sea ice in the northern sea route studies and applications*. Springer and Praxis Publishing, Chichester, UK

Kerman BR (1999) Information states in radar imagery of sea ice. *IEEE Trans Geosci Remote Sens* 37(3):1435–1446

Kotz S, Balakrishnan N, Johnson NL (2000) *Continuous multivariate distributions: models and applications*, 2nd edn. Wiley, New York

- Lantuejoul C (2002) Geostatistical simulation models and algorithms. Springer, Berlin
- Rivoirard J (1994) Introduction to Disjunctive Kriging and Non-linear Geostatistics. Clarendon Press (Oxford)
- Schabenberger O, Gotway CA (2005) Statistical Methods for Spatial Data Analysis. Chapman and Hall/CRC
- Wackernagel H (2003) Multivariate geostatistics: an introduction with application. Springer, Berlin

Index

A

Abu Dhabi, 43–52
Accuracy Assessment, 20, 37, 47–48, 177, 182–183
Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), 178–181
AlSat-1, 12
Apalachicola Bay, 69–77
Archival satellite imagery, 174
Aridzone, 17–25
Atmosphere modelling, 160
Atmospheric correction, 56–57, 59–61, 72, 74

B

Baltic Sea, 55–66, 72
BILSAT-1, 12
BIRD, 8, 204
Body pointing, 13

C

Camera, 7–11, 79, 174
Cartography, 9, 13
Case-2 water algorithms, 57
C-band, 239–252
Change analysis, 178, 257
Change detection, 18–19, 21, 159–170, 177, 180–181, 183, 256–257, 269–279
CHRIS, 8, 55
Climate change, 97–98, 109, 126, 130–131, 135, 159–170, 188, 215, 218
CNES, 2, 10
Concentration maps, 72
Confusion Matrix, 47, 86
Constellation, 1, 3, 7, 9–14, 141, 164–165
Correlation coefficient, 103–104, 191, 272
Co-seismic deformation field, 240, 242
COSMOS, 12
Cost-effective, 2–5

Coverage, 1, 10, 12–13, 29, 33–34, 36, 71–72, 98, 115, 139, 146–147, 169, 174, 189, 202, 218, 220, 227, 240, 245
Cryosphere, 79, 84–88, 132

D

DEIMOS-1, 12–13
DEM, 13, 115, 118, 179, 225–226, 244–245
DInSAR, 240–248, 250–251
Dissolved organic matter, 57, 72–73, 75
DLR, 8, 14, 60, 64
DMC, 9–13

E

Early warning system, 136, 141, 160, 162–164, 202, 217–236
EPOS-RT, 165–166
EROS, 8–10, 178–179, 217, 220
ESA, 2, 8, 45, 64, 211, 240, 252
Estuary, 69, 72–73
Evapotranspiration, 202, 204

F

Field Surveys, 182–183, 213
Food security, 201–215, 234
Forest fires, 8

G

Geodesy, 125–130, 132, 134, 137, 139, 141, 147, 154, 155, 252
Geodetic Reference Frame, 127–129
Geohazards, 125–127, 135
Georeferencing, 179
GIS, 13, 20, 28–30, 40, 64, 189–190, 219, 256, 260
Global change, 1–15, 17–18, 27–28, 125–142
GMES, 2, 56, 64, 201
GMES water quality, 55, 64
GNSS, 9, 14, 127, 130–135, 139, 141, 159–170

Google Earth, 20, 182–183
 Ground sampling distance, 6–8, 11
 GSD, 6–8, 10, 12

H

Hardware development, 170
 Heavy rainfall, 217, 219, 234–235
 HJ-1, 9–11
 Hot spots, 18, 188, 207
 Human factors, 28, 187–198
 Hydrological analysis, 219, 222–224, 231, 233
 Hyper-spectral, 7–9, 11

I

Image-differencing Change Detection, 181
 Image Mosaicking, 179
 Image Transformation, 177, 180–181
 Imaging spectrometer, 7, 14, 55, 60, 72
 Interannual variability, 97–103, 105, 107–108

L

Lag time, 219, 226
 Land cover, 9, 14, 17–25, 28, 45, 47, 49–51
 Land cover change modeling, 18, 21, 28, 256–258, 260, 267
 Land reclamation, 44, 46
 Landsat, 12, 19, 29, 32–34, 44–48, 73, 175, 178–179, 210–211, 241, 246
 Landscape matrix, 18–19, 21
 Landslide
 Distribution, 174, 177, 181, 183
 mapping, 173–184
 Land use, 7–8, 27–41, 70, 138, 187–188, 256–261, 264–265
 L-band, 79–94, 239–252
 Logistic regression, 27–41, 257–259, 262–263

M

Maximum likelihood, 20, 46, 86, 259–260
 MERIS, 55–66, 72–74, 211
 Microwave remote sensing, 112
 Mingjiang River, 183
 Miniaturization, 2, 4, 5
 Modal Filtering, 181
 Model-based inversion, 58
 MODIS, 71–76, 111–123, 188–189, 203, 211
 Multi-spectral, 7, 9–11, 14, 45

N

Natural hazards monitoring, 126, 159–170, 234
 NigeriaSat-1, 12

Normalized Difference Vegetation Index (NDVI), 46, 180–181, 203, 205, 207, 209–212, 214–215

O

Optically complex waters, 57–58
 Orbital plane, 13

P

Panchromatic, 6–7, 11
 Phytoplankton, 7–8, 56–57, 71, 75
 Polarimetry, 80
 Population Density, 189–190, 192–193, 196
 Prediction, 13, 30–32, 36, 38–39, 44, 100, 126, 138–139, 167, 258–259, 262, 264–265, 267
 Principal component inversion algorithm, 56, 59–60, 64
 PROBA, 8

R

RADAR, 9, 79–80, 89, 132–133, 135, 139, 163–164, 174, 179, 240, 243–246, 252, 270
 Radiometric Normalization, 179–180
 Rapid development, 27, 43
 RapidEye, 9–10, 12–14
 Relative Operating Characteristic, 31
 Remote Sensing, 2, 5, 8, 10, 14, 18, 28, 43–52, 55–66, 69–77, 79, 84, 112–114, 135, 169, 174, 183, 201–215, 240, 252, 256, 270
 Revisit time, 6–8, 13
 RISAT, 9
 River basins, 21, 219, 225–226

S

Satellite data, 5, 112, 188
 ScanSAR, 240–241, 246–247
 Sea ice, 146, 163, 269–279
 Sea level change, 131, 135–136, 145–155, 163
 Sichuan Earthquake, 175–178, 181, 183
 Small satellite missions, 1–15
 Snow Cover, 79–84, 88, 90–92, 97–109, 112–123
 Snowpack, 82, 89–92, 106, 112, 115–123
 Soil moisture, 89, 91–92, 93, 163
 Space
 geodesy, 127, 130, 137, 154
 technology, 2, 4–5
 Spatial Logistic Regression, 27–41
 Spatial pattern analysis, 17–25

- Spatial resolution, 5, 7–8, 20–21, 44, 46,
71–72, 74, 98, 100, 112–113, 139, 174,
180, 202, 219, 232, 275
- Spectral resolution, 5–8, 55, 57–58
- Spectrometer, 7, 9, 55, 60, 72
- SPOT-Vegetation, 203, 205, 210–211
- Spring snowmelt dates, 97–109
- Suspended matter, 65, 72
- Swath width, 7, 12, 14, 178
- Synthetic Aperture Radar (SAR), 79, 270
- T**
- TARANIS, 10
- Temporal resolution, 1, 5, 7, 9, 13, 98, 113,
127, 129, 131, 139, 141, 164, 168, 219
- TES, 8
- Thematic Mapper (TM), 19, 29, 33, 175,
178–181
- Tibetan Plateau, 111–123, 175
- TOPSAT, 10–11
- U**
- UAE, 44, 47
- UK-DMC-1, 12
- V**
- Variogram, regionalization model, 269–279
- Vegetation condition, 203, 205
- Vertical tectonic motions around the Atlantic
Ocean, 146–147
- W**
- Water
- cycle, 18, 125, 127, 132, 135, 137–138
- quality, 55–66, 69–77
- Wenchuan, 160, 175, 179, 239–252