Frontier Technologies for Infrastructures Engineering

Editors:
Shi-Shuenn Chen & Alfredo H-S. Ang

CRC Press
Taylor & Francis Group
A BALKEMA BOOK

Structures and Infrastructures 4

Frontier Technologies for Infrastructures Engineering

Structures and Infrastructures Series

ISSN 1747-7735

Book Series Editor:

Dan M. Frangopol
Professor of Civil Engineering and
Fazlur R. Khan Endowed Chair of Structural Engineering and Architecture
Department of Civil and Environmental Engineering
Center for Advanced Technology for Large Structural Systems (ATLSS Center)
Lehigh University
Bethlehem, PA, USA

Volume 4

Frontier Technologies for Infrastructures Engineering

Edited by

Shi-Shuenn Chen[1] & Alfredo H-S. Ang[2]

[1]National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan
[2]University of California, Irvine, USA
Taiwan Building Technology Center (NTUST), Taipei, Taiwan

CRC Press
Taylor & Francis Group
Boca Raton London New York Leiden

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A BALKEMA BOOK

Colophon

Book Series Editor:
Dan M. Frangopol

Volume Editors:
Shi-Shuenn Chen & Alfredo H-S. Ang

Cover illustration:
Cable Stayed Twin Bridge at Malpensa Airport (1998), Milan, Italy. Designed by Professor Francesco Martinez y Cabrera. Managed by SEA Aeroporti Company.

Taylor & Francis is an imprint of the Taylor & Francis Group, an informa business

Typeset by Charon Tec Ltd (A Macmillan company), Chennai, India
Printed and bound in Great Britain by Antony Rowe (a CPI Group company), Chippenham, Wiltshire

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

International Workshop on Frontier Technologies for Infrastructures Engineering (2008 : Taipei, Taiwan)
Frontier technologies for infrastructures engineering / edited by Shi-Shuenn Chen & Alfredo H-S. Ang.
p. cm. — (Structures and infrastructures series, ISSN 1747-7735 ; v. 4)
Papers of the International Workshop on Frontier Technologies for Infrastructures Engineering held in Taipei, Taiwan, Oct. 23–25, 2008.
Includes bibliographical references and index.
ISBN 978-0-415-49875-3 (hardcover : alk. paper) — ISBN 978-0-203-87559-9 (e-book)
1. Civil engineering—Technological innovations—Congresses. 2. Public works—Technological innovations—Congresses. I. Chen, Shi-Shuenn. II. Ang, Alfredo Hua-Sing, 1930– III. Title. IV. Series.
TA5.I745 2008
624—dc22

2009007325

Published by: CRC Press/Balkema
P.O. Box 447, 2300 AK Leiden, The Netherlands
e-mail: Pub.NL@taylorandfrancis.com
www.crcpress.com – www.taylorandfrancis.co.uk – www.balkema.nl

ISBN13 978-0-415-49875-3(Hbk)
ISBN13 978-0-203-87559-9(eBook)
Structures and Infrastructures Series: ISSN 1747-7735
Volume 4

Table of Contents

Editorial

Welcome to the New Book Series *Structures and Infrastructures*.

Our knowledge to model, analyze, design, maintain, manage and predict the life-cycle performance of structures and infrastructures is continually growing. However, the complexity of these systems continues to increase and an integrated approach is necessary to understand the effect of technological, environmental, economical, social and political interactions on the life-cycle performance of engineering structures and infrastructures. In order to accomplish this, methods have to be developed to systematically analyze structure and infrastructure systems, and models have to be formulated for evaluating and comparing the risks and benefits associated with various alternatives. We must maximize the life-cycle benefits of these systems to serve the needs of our society by selecting the best balance of the safety, economy and sustainability requirements despite imperfect information and knowledge.

In recognition of the need for such methods and models, the aim of this Book Series is to present research, developments, and applications written by experts on the most advanced technologies for analyzing, predicting and optimizing the performance of structures and infrastructures such as buildings, bridges, dams, underground construction, offshore platforms, pipelines, naval vessels, ocean structures, nuclear power plants, and also airplanes, aerospace and automotive structures.

The scope of this Book Series covers the entire spectrum of structures and infrastructures. Thus it includes, but is not restricted to, mathematical modeling, computer and experimental methods, practical applications in the areas of assessment and evaluation, construction and design for durability, decision making, deterioration modeling and aging, failure analysis, field testing, structural health monitoring, financial planning, inspection and diagnostics, life-cycle analysis and prediction, loads, maintenance strategies, management systems, nondestructive testing, optimization of maintenance and management, specifications and codes, structural safety and reliability, system analysis, time-dependent performance, rehabilitation, repair, replacement, reliability and risk management, service life prediction, strengthening and whole life costing.

This Book Series is intended for an audience of researchers, practitioners, and students world-wide with a background in civil, aerospace, mechanical, marine and automotive engineering, as well as people working in infrastructure maintenance, monitoring, management and cost analysis of structures and infrastructures. Some volumes are monographs defining the current state of the art and/or practice in the field, and some are textbooks to be used in undergraduate (mostly seniors), graduate and

postgraduate courses. This Book Series is affiliated to *Structure and Infrastructure Engineering* (http://www.informaworld.com/sie), an international peer-reviewed journal which is included in the Science Citation Index.

It is now up to you, authors, editors, and readers, to make *Structures and Infrastructures* a success.

Dan M. Frangopol
Book Series Editor

About the Book Series Editor

Dr. Dan M. Frangopol is the first holder of the Fazlur R. Khan Endowed Chair of Structural Engineering and Architecture at Lehigh University, Bethlehem, Pennsylvania, USA, and a Professor in the Department of Civil and Environmental Engineering at Lehigh University. He is also an Emeritus Professor of Civil Engineering at the University of Colorado at Boulder, USA, where he taught for more than two decades (1983–2006). Before joining the University of Colorado, he worked for four years (1979–1983) in structural design with A. Lipski Consulting Engineers in Brussels, Belgium. In 1976, he received his doctorate in Applied Sciences from the University of Liège, Belgium, and holds two honorary doctorates (Doctor Honoris Causa) from the Technical University of Civil Engineering in Bucharest, Romania, and the University of Liège, Belgium. He is a Fellow of the American Society of Civil Engineers (ASCE), American Concrete Institute (ACI), and International Association for Bridge and Structural Engineering (IABSE). He is also an Honorary Member of both the Romanian Academy of Technical Sciences and the Portuguese Association for Bridge Maintenance and Safety. He is the initiator and organizer of the Fazlur R. Khan Lecture Series (www.lehigh.edu/frkseries) at Lehigh University.

Dan Frangopol is an experienced researcher and consultant to industry and government agencies, both nationally and abroad. His main areas of expertise are structural reliability, structural optimization, bridge engineering, and life-cycle analysis, design, maintenance, monitoring, and management of structures and infrastructures. He is the Founding President of the International Association for Bridge Maintenance and Safety (IABMAS, www.iabmas.org) and of the International Association for Life-Cycle Civil Engineering (IALCCE, www.ialcce.org), and Past Director of the Consortium on Advanced Life-Cycle Engineering for Sustainable Civil Environments (COALESCE). He is also the Chair of the Executive Board of the International Association for Structural Safety and Reliability (IASSAR, www.columbia.edu/cu/civileng/iassar) and the Vice-President of the International Society for Health Monitoring of Intelligent Infrastructures (ISHMII, www.ishmii.org). Dan Frangopol is the recipient of several prestigious awards including the 2008 IALCCE Senior Award, the 2007 ASCE Ernest Howard Award, the 2006 IABSE OPAC Award, the 2006 Elsevier Munro Prize, the 2006 T. Y. Lin Medal, the 2005 ASCE Nathan M. Newmark Medal, the 2004 Kajima

Research Award, the 2003 ASCE Moisseiff Award, the 2002 JSPS Fellowship Award for Research in Japan, the 2001 ASCE J. James R. Croes Medal, the 2001 IASSAR Research Prize, the 1998 and 2004 ASCE State-of-the-Art of Civil Engineering Award, and the 1996 Distinguished Probabilistic Methods Educator Award of the Society of Automotive Engineers (SAE).

Dan Frangopol is the Founding Editor-in-Chief of *Structure and Infrastructure Engineering* (Taylor & Francis, www.informaworld.com/sie) an international peer-reviewed journal, which is included in the Science Citation Index. This journal is dedicated to recent advances in maintenance, management, and life-cycle performance of a wide range of structures and infrastructures. He is the author or co-author of over 400 refereed publications, and co-author, editor or co-editor of more than 20 books published by ASCE, Balkema, CIMNE, CRC Press, Elsevier, McGraw-Hill, Taylor & Francis, and Thomas Telford and an editorial board member of several international journals. Additionally, he has chaired and organized several national and international structural engineering conferences and workshops. Dan Frangopol has supervised over 70 Ph.D. and M.Sc. students. Many of his former students are professors at major universities in the United States, Asia, Europe, and South America, and several are prominent in professional practice and research laboratories.

For additional information on Dan M. Frangopol's activities, please visit www.lehigh.edu/~dmf206/

Preface

This volume contains all the papers of the twenty lectures presented at the *International Workshop on Frontier Technologies for Infrastructures Engineering* that was held in Taipei, Taiwan at the Taiwan Building Technology Center (TBTC) of the National Taiwan University of Science & Technology (Taiwan Tech) on 23–25 October 2008. Each of the papers highlights the frontier or emerging technology that can serve to improve the practice of infrastructure engineering. The papers are intended to describe what or how improvements are possible through implementation of the respective topical technologies.

The Workshop was organized for the purpose of introducing some of the technologies that have been developed in the last two decades that will be of benefit to the design, construction, and maintenance of civil infrastructure systems. The authors/lecturers of the papers in the Workshop were invited based on the records of accomplishments in their respective fields of expertise, and all are world renowned for their work. The publication of all the papers in this single volume should benefit also the practicing engineers who were unable to attend the Workshop.

It is often said that there is a major gap between engineering research and professional practice. Research, of course, is necessary as the first step in advancing the technology of engineering practice. The Workshop in October 2008 and the publication of this volume were intended to help fill this gap, by bringing some of the recent results of research to the attention of the practicing community of infrastructure engineers. The topics covered in the volume can be gleaned from the titles in the Table of Contents, which specifically include the following:

- Special issues in the dynamics of buildings.
- Durability of infrastructures through fatigue and fracture resistance.
- Corrosion in reinforced concrete construction
- New methods and state-of-the-art of optimal design of structures.
- Reliability and reliability-based design of structures.
- Risk assessment and risk management.
- Life-cycle performance and cost in civil engineering.
- Health monitoring for maintenance of infrastructures
- Applications in geotechnical and earthquake engineering
- Innovations in construction and project management

Each of the papers, or several papers, can only give a brief summary of the pertinent topic, and therefore can serve only to provide introductions to the respective topics. Through the references cited in the papers, more complete treatment of each topic may be found.

The Workshop and the publication of this volume were supported by the Taiwan Tech. Our sincere thanks to the faculty and administrative staff of the TBTC for their cooperation and assistance in the preparation of the event. Finally, our thanks to all the lecturers for their superb presentations at the Workshop, and for preparing their papers in this volume. It is our hope that the Workshop and the publication of this volume will serve to advance the technologies for infrastructures engineering.

Shi-Shuenn Chen and A. H-S. Ang

Organization

Organized by

National Taiwan University of Science & Technology (Taiwan Tech)
Taiwan Building Technology Center (TBTC)
Department of Construction Engineering, Taiwan Tech

In Collaboration with

National Science Council, Taiwan
Water Resources Agency, Taiwan
Ministry of Transportation and Communications, Taiwan
Taiwan Institute of Steel Construction, Taiwan
American Society of Civil Engineers, USA
Chinese Institute of Civil and Hydraulic Engineering, Taiwan
International Association for Bridge Management and Safety
International Association for Life-Cycle Civil Engineering
International Association for Structural Safety and Reliability

Sponsors

Architecture and Building Research Institute
Bureau of High Speed Rail, MOTC
CECI Engineering Consultants, Inc.
China Steel Corporation
Chinese Society of Pavement Engineering
Construction and Planning Agency
Continental Engineering Corporation
CTCI Corporation
Department of Rapid Transit Systems, TCG
Directorate General of Highways, MOTC
Land Engineering Consultants Co., Ltd.
Moh and Associates, Inc.
Public Construction Commission, Executive Yuan
RSEA Engineering Corporation
Ruentex Group

Sinotech Engineering Consultants, Ltd.
Taipei Professional Civil Engineers Association
Taipei Professional Geotechnical Engineers Association
Taipei Professional Structural Engineers Association
Taipei Rapid Transit Corporation
Taiwan Area National Expressway Engineering Bureau
Taiwan Area National Freeway Bureau
Taiwan Asphalt Industrial Association
Taiwan Giken Seisakusho Corporation, Ltd.
Taiwan High Speed Rail Corporation
Taiwan Professional Civil Engineers Association
Taiwan Professional Geotechnical Engineers Association
Taiwan Professional Structural Engineers Association
Tung Ho Steel Enterprise Corporation

Committees

Organizing Committee

Chair – Chen, Shi-Shuenn (Taiwan Tech)
Co-Chair – Ang, Alfredo H-S. (U.C. Irvine)
Members – Chen, Cheng-Cheng (Taiwan Tech)
Chen, Chun-Tao (Taiwan Tech)
Chen, Hung-Ming (Taiwan Tech)
Cherng, Rwey-Hua (Taiwan Tech)
Ching, Jianye (National Taiwan University)
Chou, Jui-Sheng (Taiwan Tech)
Hsieh, Yo-Ming (Taiwan Tech)
Lee, Wei-Feng (Taiwan Tech)
Leu, Sou-Sen (Taiwan Tech)
Liao, Kuo-Wei (Tamkang University)
Ou, Chang-Yu (Taiwan Tech)
Ou, Yu-Chen (Taiwan Tech)
Yang, I-Tung (Taiwan Tech)
Young, Chin-Huai (Taiwan Tech)

Advisory Committee

Chair – Ou, Chin-Der (Taiwan High Speed Rail Corporation)
Members – Chang, Chi-Te (Taiwan Dept. of Rapid Transit Systems)
Chang, Kuo-Chun (National Taiwan University)
Chang, Shuenn-Yih (National Taipei Univ. of Technology)
Chen, How-Ji (National Chung-Hsing University)
Chen, Jin-Yuan (Directorate General of Highways)
Chen, Shen-Hsien (Water Resources Agency)
Chen, Tung-Yang (National Cheng-Kung University)
Cho, Jui-Nian (Sinotech Engineering Consultants)
Ho, Ming-Chin (Architecture and Building Research Institute)
Hwang, Shyh-Jian (Chinese Soc. of Structural Engineering)
Li, Chien-Chung (CECI Engineering Consultants)
Li, Thay-Ming (Taiwan Area National Freeway Bureau)
Moh, Za-Chieh (Moh and Associates)

Pang, Jar-Hua (Bureau of High Speed Rail)
Shue, Chyi (National Kaohsiung Univ. of App. Sciences)
Tien, Yong-Ming (National Central University)
Tseng, Dar-Jen (Taiwan Area National Expressway Eng. Bureau)
Yang, Yeong-Bin (Chinese Inst. of Civil and Hydraulic Eng.)
Yeh, Keh-Chia (National Chiao-Tung University)
Yeh, Shih-Wen (Construction and Planning Agency)
Yin, Yen-Liang (Ruentex Group)
Yu, John T. (CTCI Corporation)

About the Editors

Shi-Shuenn Chen – Dr. Chen, who received his doctorate in civil engineering from the University of California, Berkeley (1987), is currently President of the National Taiwan University of Science and Technology (Taiwan Tech), where he has previously served as Vice-President, and Dean of academic affairs. He established the Taiwan Building Technology Center in Taiwan while pursuing a First-Class University to foster the development of building technologies at the cutting edge of research. He is also a full professor in the Department of Construction Engineering at NTUST (Taiwan Tech) and has previously served as Department Chair. Dr. Chen also currently serves as a member of the Public Construction Commission, under the Executive Yuan of the ROC, and of the Supervisory Board of Science Education, under the Ministry of Education. He is a Managing Director on the board of directors and a Fellow of the Chinese Institute of Civil and Hydraulic Engineering (CICHE), as well as serving as the editor-in-chief of the *Journal of the Chinese Institute of Engineers*. Dr. Chen's previous positions also include those of deputy director general in the National Expressway Engineering Bureau, President of the Chinese Association of Structure Engineers and Managing Director on the board of directors of the Association of National Universities. Honors received by Dr. Chen include the Outstanding Engineering Professor Award from the Chinese Institute of Engineers and the Excellent Research Paper Award from the CICHE as well as being recognized as an Outstanding Civil Engineering Alumnus by the National Taiwan University. Dr. Chen has published extensively in his main areas of research: earthquake engineering.

Alfredo H-S. Ang – Dr. Ang is currently Director of the Taiwan Building Technology Center at the NTUST, and Research Professor and Professor Emeritus at the University of California in Irvine, California, USA. He was on the faculty of Civil and Environmental Engineering at the University of Illinois at Urbana-Champaign in 1959–1988 where he received his Ph.D. in 1959, and is Emeritus Professor since 1988. He is a member of the US National Academy of Engineering (elected in 1976), and Honorary Member of the ASCE (1991). His main area of research is on the application of probability and reliability in civil and structural engineering, with emphasis on safety of engineering systems,

including seismic risk and earthquake engineering, quantitative risk assessment (QRA) and risk management, and life-cycle cost consideration. He has published over 400 papers and articles, and also a two-volume textbook on *probability concepts in engineering,* which have been translated into several languages; the 2nd edition of Vol I was published in February 2006. During his academic career, he has directed 55 Ph.D. students and countless postdoctoral researchers from many parts of the world. He has given keynote papers in numerous major national and international conferences. During his career, he has been consultant and technical adviser to government and industry, and served on a number of advisory boards, both in the U.S. and abroad. He was the International Director on the Board of Directors of ASCE in 1998–2001, and is currently the ASCE representative to the Asian Civil Engineering Coordinating Council (ACECC). He is also a Fellow of the ASME, Associate Fellow of the AIAA, and a founding member of IASSAR, Honorary President of IALCCE (International Association of Life-Cycle Civil Engineering). Among his honors and awards are the ASCE N.M. Newmark Medal, the A. Freudenthal Medal, E. Howard Award, Walter Huber Research Prize, State-of-Art Award; Senior Research Award from the ASEE and Research Award from IASSAR; Distinguished Research Award from the University of California, Irvine; Distinguished Engineering Alumni Award from the University of Illinois at Urbana-Champaign; and the 2005 International Prize from the Japan Society of Civil Engineers.

Chapter 1

Innovative conceptual design of vibration sensitive buildings

S.L. Lee, C.G. Koh & D.K.H. Chua

Department of Civil Engineering, National University of Singapore, Singapore

ABSTRACT: In the semiconductor and precision engineering industry, the production buildings normally have to conform to very stringent vibration criteria. Allowable velocity specification for the structural floor is typically defined in the order of micrometers per second. The structural design of these micro-vibration sensitive poses a big challenge especially in busy metropolitan cities. In this paper, an innovative approach of isolation-and-shield is presented to mitigate vibrations from multiple sources such as traffic, wind and machineries. The approach has proven to be effective and robust in more than a dozen of actual vibration sensitive buildings the authors have advised upon. Practical guideline on the design of the floor system against human footfall in terms of a simple stiffness-frequency index is presented. The index has been established based on finite element analysis and correlated with field verification measurements.

1 Introduction

High-end electronic components require increasing precision optical and micro-assembly equipment in semiconductor wafer fabrication plants. Besides, the production and quality control processes are extremely sensitive to vibrations so that the structure has to be carefully designed to avoid unnecessary disruption due to vibration far below human perceptibility. Due to the use of increasingly sophisticated manufacturing processes, the vibration requirements are likely to become more stringent (Gordon 1991; Leung and Papadimos 1999) and have become a critical factor in the design of such facilities. This has direct effect on the rejection rate of products, which affects not only directly on profitability but also the reputation of manufacturer. Presently, this strict requirement often translates into floor vibration specification in the range of 3.2–12.7 micrometers per second (or 125–500 micro-inches per second). In the expansion or upgrading of existing buildings, the task of achieving strict vibration requirements would become even more challenging due to site constraints and ongoing operations.

In the micro-vibration study of vibration sensitive buildings, several sources of vibration have to be considered. The first is essentially ground borne in nature comprising excitations mainly from traffic and other activities on the ground. The vibration from such source is transmitted through the soil medium and the supporting foundation system into vibration sensitive areas. The second source is wind excitation on the building. This can constitute an important consideration as the roof systems for such buildings are often long span and flexible to provide maximum production space and

layout flexibility. The third source of vibration is human footfall (walking) load in the production area. This would primarily affect the design of the floor system of vibration-sensitive areas such as the main production area of the building. Predominantly vertical, footfall-induced vibrations are dependent on the walking rate of personnel and the dynamic characteristics of the floor system (Ungar and White 1970; Ellis et al 1997). Past experience has shown that footfall could cause floor vibrations even higher than traffic-induced vibrations. Furthermore, while it is essential to provide proper damping systems for the mechanical and electrical (M&E) equipments, the vibrations caused are another source of vibration to be considered in the structural system.

2 Ground borne vibrations

There are several strategies to mitigate vibration due to ground borne excitation. One strategy is to use passive dampers such as rubber bearings. In this way a compliant link is introduced between the building and the ground, so that the predominant modes of the system are shifted down in natural frequencies, well below the frequency range of the disturbing vibrations (Northwood, 1977; Koh, 1986). This is known as base isolation which has become a viable option for ground excitations of significant level such as earthquakes. But its practicability with regards to mitigating micro-vibration of high-tech buildings is not well investigated. The second strategy is to have long trenches as wave barriers, surrounding either the source of vibration (active isolation) or the production building or area (passive isolation). Nevertheless, the trenches have to be fairly deep for this strategy to work effectively.

Another practical strategy of mitigating ground borne vibrations is providing isolation gap surrounding the production building which houses the micro-vibration sensitive area. Bachmann and Ammann (1987) proposed disconnecting structural parts so as to reduce transmission of the waves. Expanding on this idea, our recommended strategy to reduce the ground borne vibration has been the provision of a "shield building" all around the vibration sensitive production building. The shield building serves to absorb dynamic forces due to wind and other external activities, and is isolated from the production building to minimize the transmission of vibrations. Ground borne waves are dissipated and scattered by the foundation system of the shield building, which could incorporate additional piles if necessary to increase the shield effects. This strategy has proven to be effective and robust based on the authors' involvement in more than a dozen of vibration sensitive buildings in Singapore, Malaysia and China. Figures 1 and 2 show the cross section view and plan view of a typical wafer fabrication plant designed adopting the isolation-and-shield strategy.

In practice, it is necessary to include field measurement of ground motion (input) and verification measurement of response in the vibration-sensitive areas when the structure is completed. Ground motion is measured (usually prior to construction) so as to provide an assessment of the ground borne excitation. This would be used to estimate structural dynamic response and is important in deciding the overall structural system and layout. In highly urbanized cities such as Singapore and Taipei where high-tech buildings are often built within or near urban areas, traffic excitation forms the main bulk of ambience-induced vibrations. Figure 3 shows a high-tech industrial park called Woodlands Wafer Park in Singapore and its proximity to a major expressway frequented by heavy trucks, a subway line, an on-grade train railway and local access

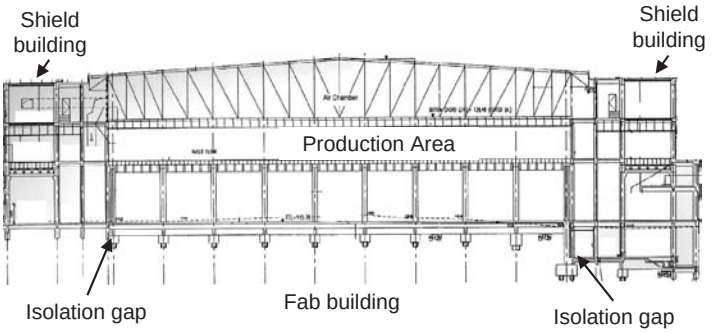


Figure 1 Cross section of a wafer fabrication factory with shield building.

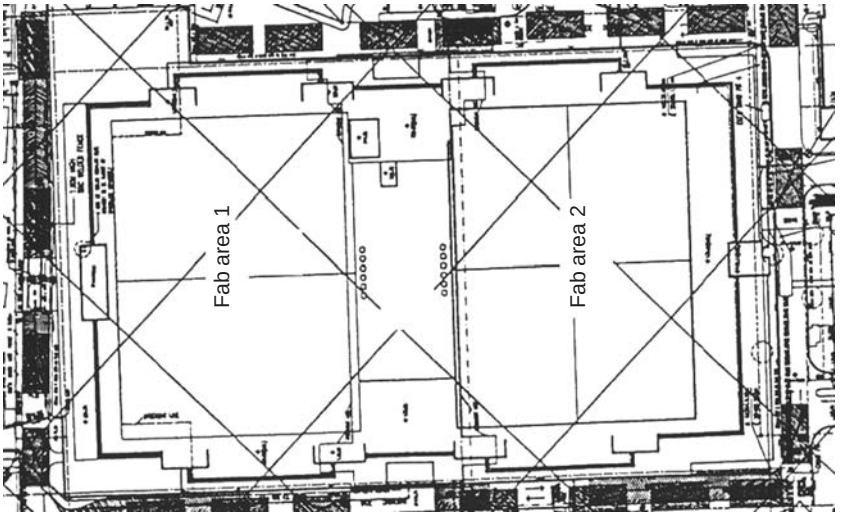


Figure 2 Plan view of a wafer fabrication factory with shield building.

roads. The presence of bus stop and traffic junction, where vehicles decelerate, stop and accelerate, could also cause undesirable vibrations to the production areas. Even in a relatively quiet site such as a new industrial park in Kuching of East Malaysia, it is recommended to account for future increase in traffic if further development of the industrial park is anticipated. In this regard, ground motion measured in the matured Woodlands Wafer Park may be used as the "design ground spectrum", as shown in Figure 4. The figure also presents the family of generic vibration criteria VC-A to E and various ISO guidelines commonly adopted in the industry.

It is imperative that ground motion measurement be carried out correctly. The measured signal would not be representative of the actual ground motion if the sensor is laid on top soil of the ground surface or attached to a concrete (or steel) block loosely inserted in the soil. Wrong choice of sensors with insufficient resolution is another common problem. In the ground motion measurement, it is advisable to have a concrete cube block of at least $1 \times 1 \times 1$ m cast-in-situ and embedded in an

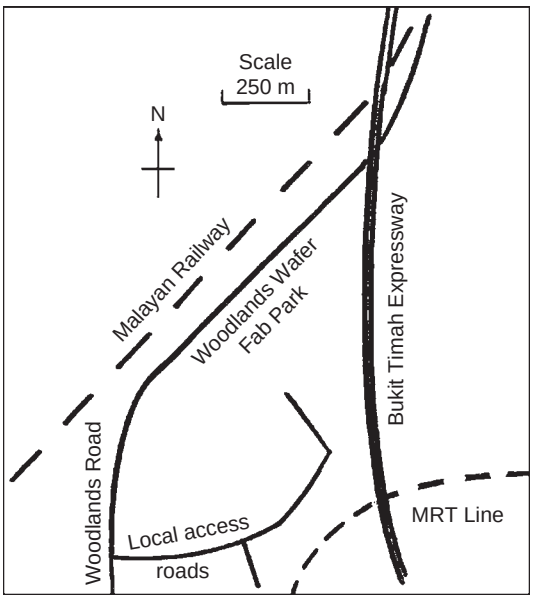


Figure 3 Woodlands wafer fab park in Singapore.

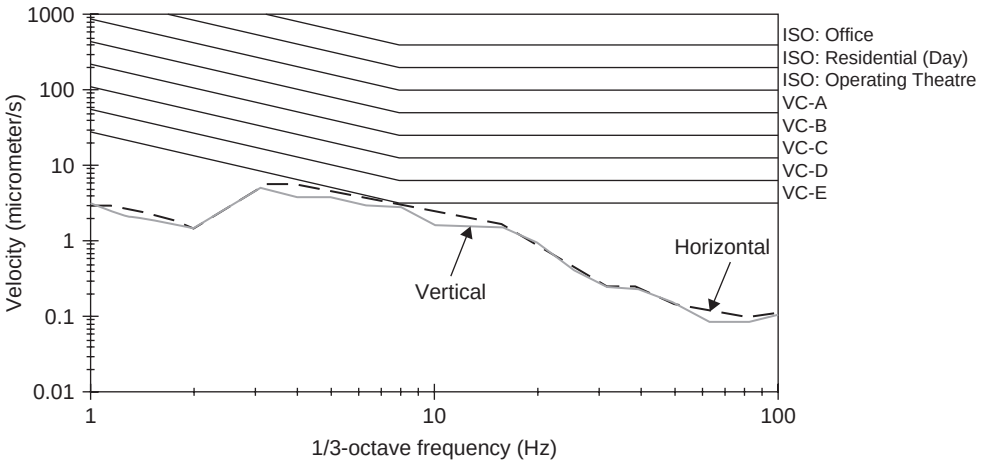


Figure 4 Woodlands velocity spectrum in comparison with industry standards.

excavated pit of sufficient depth to ensure good bonding with reasonably firm soil. Figure 5 schematically shows the measurement pit and the concrete block. Heavy mass accelerometers with high resolution are attached securely to a steel block which is glued to the top surface of the concrete block.

To reduce the effects of ground borne vibrations, the proposed isolation-and-shield strategy is highly recommended whereby the vibration-sensitive building is isolated from a shield building. A plane-strain finite element analysis is carried out to study the

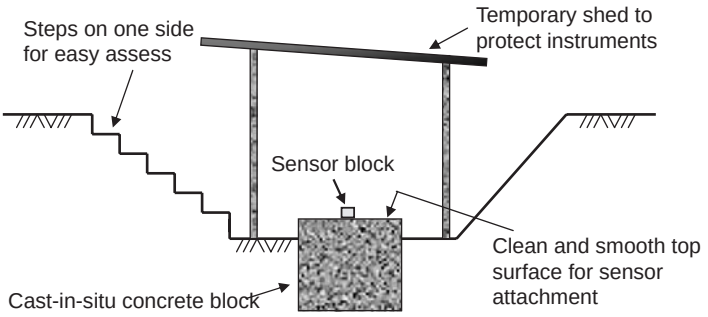


Figure 5 Measurement of ground motion.

effect of this strategy. The building modelled was a wafer fabrication plant comprising three floors with each floor about 7 m high except for the first which was 6.2 m high. The plant (including the shield building) had an overall dimension of 108 m length and 90 m width. Shear walls were provided for lateral stiffness. The vibration sensitive production area was located on the third floor. A smeared property for the piles was adopted to give a reasonable representation of the soil-structure interaction for traffic induced vibrations. Infinite elements were utilised for the lateral boundaries of the soil medium to simulate the far field effect by absorbing the approaching waves.

From random theory, the power spectrum of the structural displacement response $Y^2_{traf}(\omega)$ due to traffic at frequency ω is given by

$$Y^2_{traf}(\omega) = F^2_{traf}(\omega) \bullet H^2(\omega) \tag{1}$$

where $F^2_{traf}(\omega)$ is the power specrtrum of the traffic load and $H(\omega)$ the transfer function. $H^2(\omega)$ is derived from the dynamic stiffness S of the overall finite element assemblage obtained as:

$$[\mathrm{S}] = -\omega^2[M] + i\omega[C] + [K] \tag{2}$$

in which M, C and K are the overall mass, damping and stiffness matrices, respectively.

Figure 6 shows the horizontal velocity obtained at the production area from the above analysis in comparison with that measured on site. The analytical results are in good agreement with field measurements for frequencies below about 10 Hz. At higher frequencies, vibrations from M&E equipments are not accounted for in the finite element analysis and, if significant (normally in frequency range higher than 10 Hz), have to be reduced by using proper isolator or damper support system. The effectiveness of the isolation gap is illustrated in Figure 7. It is evident that the horizontal vibration levels after the gap drop by about 10 times in the frequency range of interest. At the level of the production area, the effect is smaller as depicted in Figure 8. The horizontal vibration level for the design with the isolation gap is on the average about 2 times lower than that without the isolation gap in the frequency range of interest (up to 10 Hz).

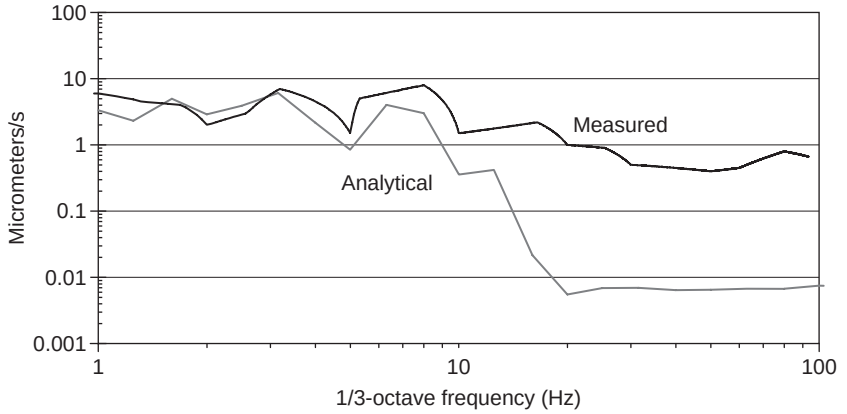


Figure 6 Analytical and measured results of horizontal velocity in production area (M&E excitations not included in analysis).

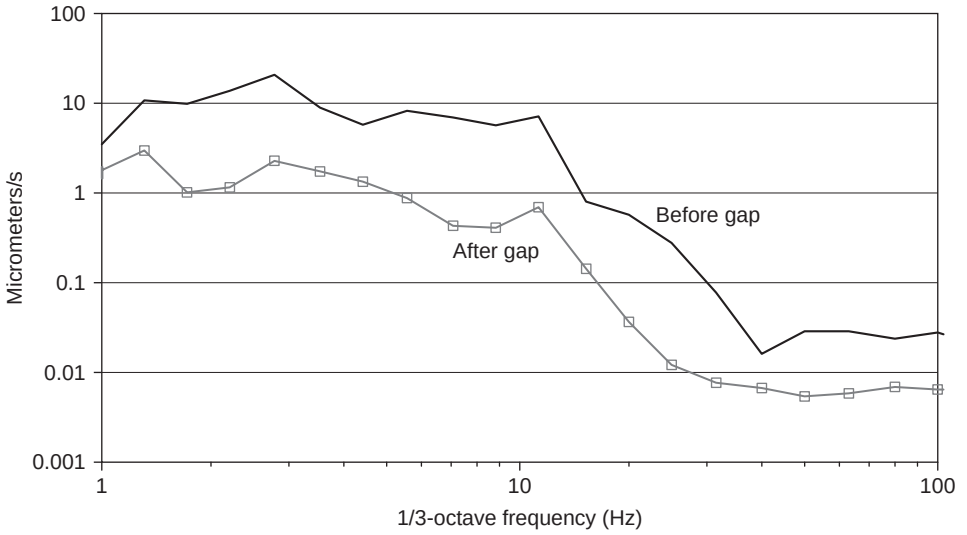


Figure 7 Horizontal vibration responses at ground floor level before and after isolation gap.

To cater for M&E requirements such as clean air ducts and chemical pipelines, high-tech factories usually require high storey height (commonly 5–7 m per storey). Coupled with the relatively heavy mass of strong floor system and high live load, the structure would be laterally flexible unless stiffened by shear walls. Asymmetry of shear wall layout should be minimized as far as possible to reduce torsional response. To facilitate the many M&E service ducts which are inevitably needed in high-tech building, shear walls of partial height are often used to provide clear opening above the wall as illustrated in Figure 9. This would minimize the effect of any change in M&E requirements on structural design and construction.

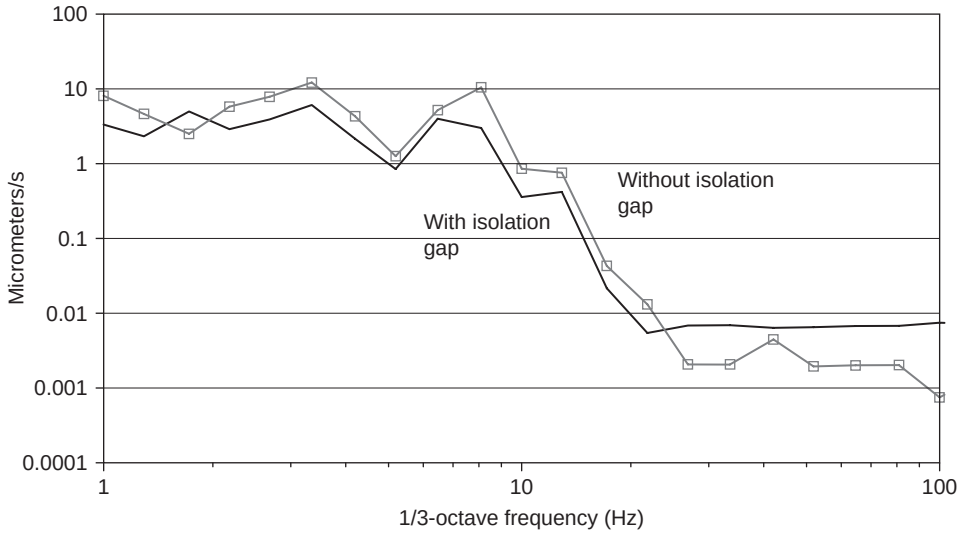


Figure 8 Horizontal velocity responses in production area with and without isolation gap.

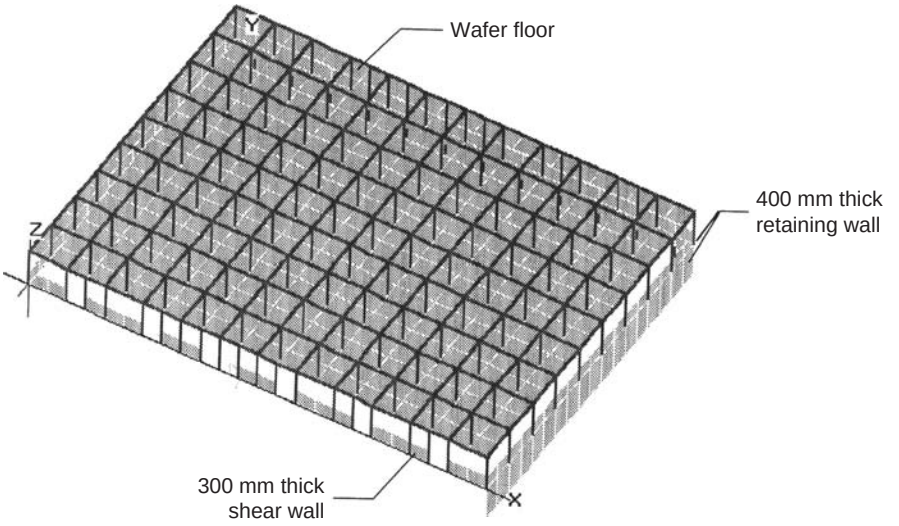


Figure 9 3D finite element model of a wafer fabrication plant.

After the structural work of the production building is completed, it is a normal practice to measure the response for verification purpose. Figure 10 shows the accelerometers and instruments used in typical response measurement on a waffle floor beam. The vibration velocities were measured at the production area level of an operational wafer fabrication building – one point in the isolated production area with VC-D criterion and the other in the shield building. It can be seen in Figure 11 that the

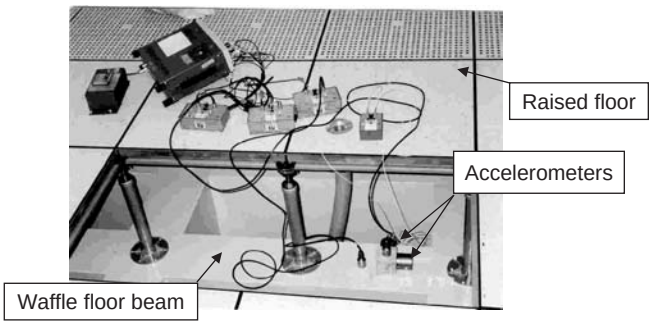


Figure 10 Vibration measurement on a waffle floor beam in a production area.

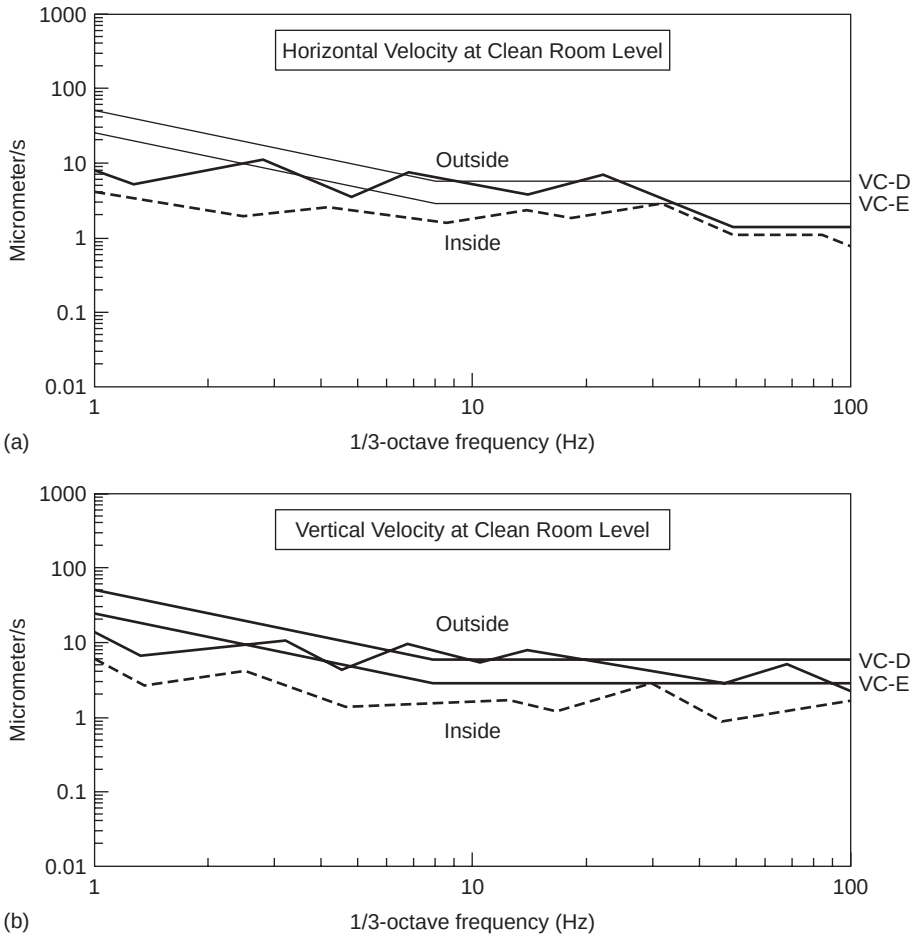


Figure 11 Measured vibration response at the clean room level inside and outside the isolated production building.

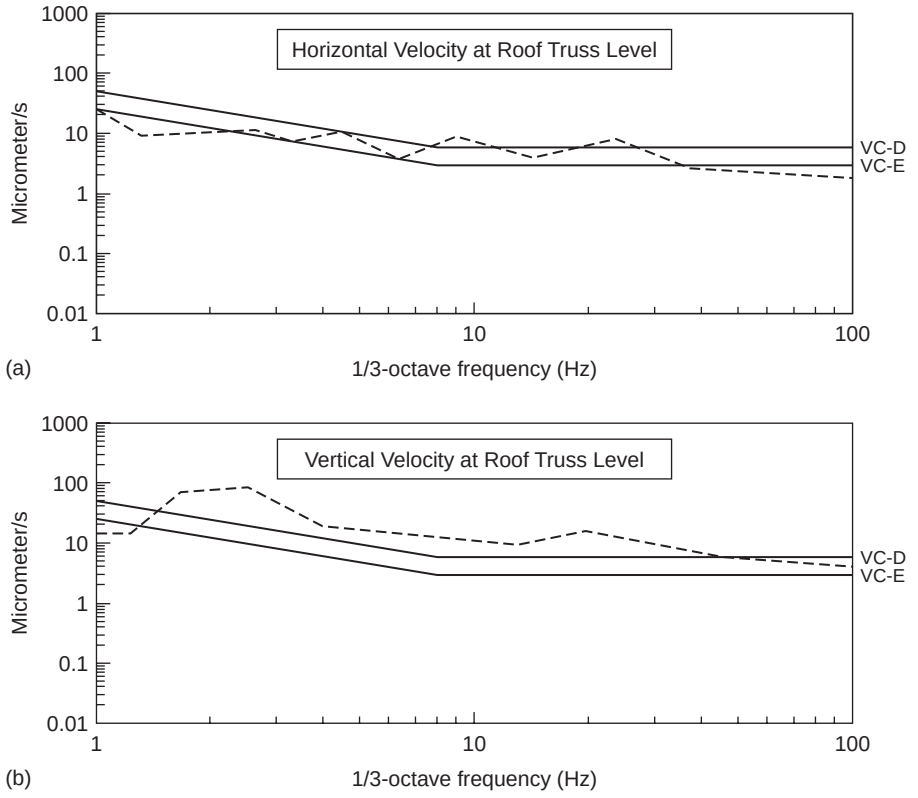


Figure 12 Measured vibration response at roof truss (supported by shield building).

vibration level within the isolated area is significantly lower than (by 5 to 10 times) than in the shield building, demonstrating the effectiveness of the isolation-and-shield strategy adopted.

3 Wind induced vibrations

The isolation-and-shield strategy also serves to directly mitigate the effects of winds on the production building, particularly vibrations caused to the roof structure which is normally long span and thus susceptible to wind excitation. Instead of stiffening the roof structure, a more cost effective way is to isolate the production building from the roof and its supporting structures. This innovative strategy ensures that not only wind excitations do not disturb the production building, but also M&E vibration sources (e.g. make-up air units) supported on the roof structure and its supporting building are isolated from the vibration sensitive production areas. Figure 12 shows the high vibrations measured at the roof truss level for a wafer fabrication building. The peak velocity in the vertical direction exceeds the VC-D criterion by about four times. In spite of the high vibrations measured at the roof truss level as shown in Figure 12, the vibration criterion is satisfied in the isolated area comfortably (as previously shown in

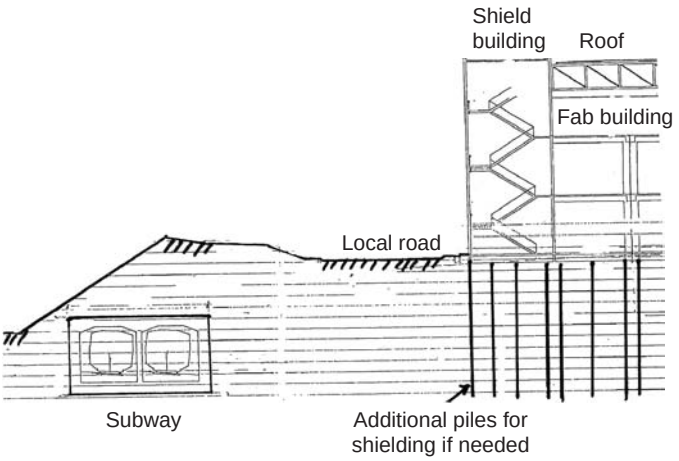


Figure 13 Isolation-and-shield concept applied to mitigate external vibrations (traffic and wind).

Figure 11). The effectiveness of this isolation-and-shield strategy is again manifested for a different source of vibration. In cases where intermediate supports such as stair cores within the production building are needed to reduce the roof span and cost, the intermediate supports and their foundations have to be isolated from the production building.

3.1 *M&E vibration sources*

Besides acting as an isolation-and-shield for the production area, the shield building provides corridor access for movement of humans and materials, as well as space for locating any vibration-causing equipment such as make-up air units. To prevent direct transmission onto the floor slab, vibration-causing equipment are put on active or passive damper systems. These damping systems, however, deteriorate over time so that the isolation-and-shield strategy would offer greater robustness for mitigating vibration in the production area.

This strategy is even more crucial when the vibration sensitive area is on the ground floor because of the direct wave transmission to the floor slab from ground borne vibration sources. If necessary, additional piles acting as a shield may be installed on the outer perimeter of the shield building (see Figure 13). These shield piles, together with the pile caps, would help absorb and scatter the incoming vibration waves. This concept has been implemented for an existing building where the production was upgraded for vibration sensitive wafer production. The VC-D production area was located on the ground floor. Due to existing constraints, the chillers system was located in the adjoining building as shown in Figure 14. To achieve isolation-and-shielding for the clean room, the ground floor slab of the production area was supported on its own micro-piles, structurally independent of the ground beams of the existing foundation. In addition, a line of contiguous micro-piles with a thick pile capping was installed along the boundary of the production area adjoining the building housing the chillers

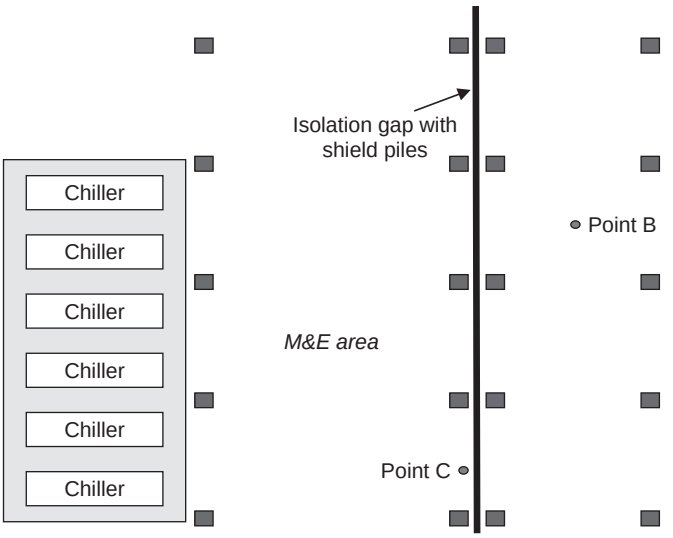


Figure 14 Isolation-and-shield concept applied to mitigate M&E vibrations.

system to provide the shield. Verification measurement as depicted in Figure 15 shows a significant reduction in vibration level. The vibration level at Point B in the clean room is 8 to 10 times lower than that measured at Point C (in the chillers building), about 20 m from the chillers. During the measurement, only one unit was in operation. When all six units are operational the vibration level in the clean room area would be expected to be raised by between 2.5 times. This would still meet the VC-D criterion, thus demonstrating the effectiveness of the isolation-and-shield system proposed.

4 Footfall induced vibrations

The footfall is another important source of vibration coming internally from operators working within the production areas. Past experience has shown that, for high-tech buildings with stringent vibration requirement, footfall can cause floor vibrations that are as significant as, if not more than, traffic-induced vibrations. To minimize the adverse effect of walking, a raised floor system is set up on top of the floor slab for the operators to walk on so that the walking load does not act directly on the floor slab but is transmitted through the raised floor system. The raised floor system is designed so that each panel acts independently to loads acting on it with little transfer of energy to the surrounding panels. The damping within the panels help to reduce the amount of vibration transferred to the floor slab. To provide further reduction in vibration, the vibration sensitive equipment are set on special damped mountings which rest on the floor slab.

The floor system commonly employed in the design of production area is a waffle slab system. Essentially, it is an assemblage of evenly spaced deep ribs supported on main beams with typical plan and section as shown in Figure 16. The ribs are spaced at a distance to accommodate the vertical pedestals of the raised floor. The main beams

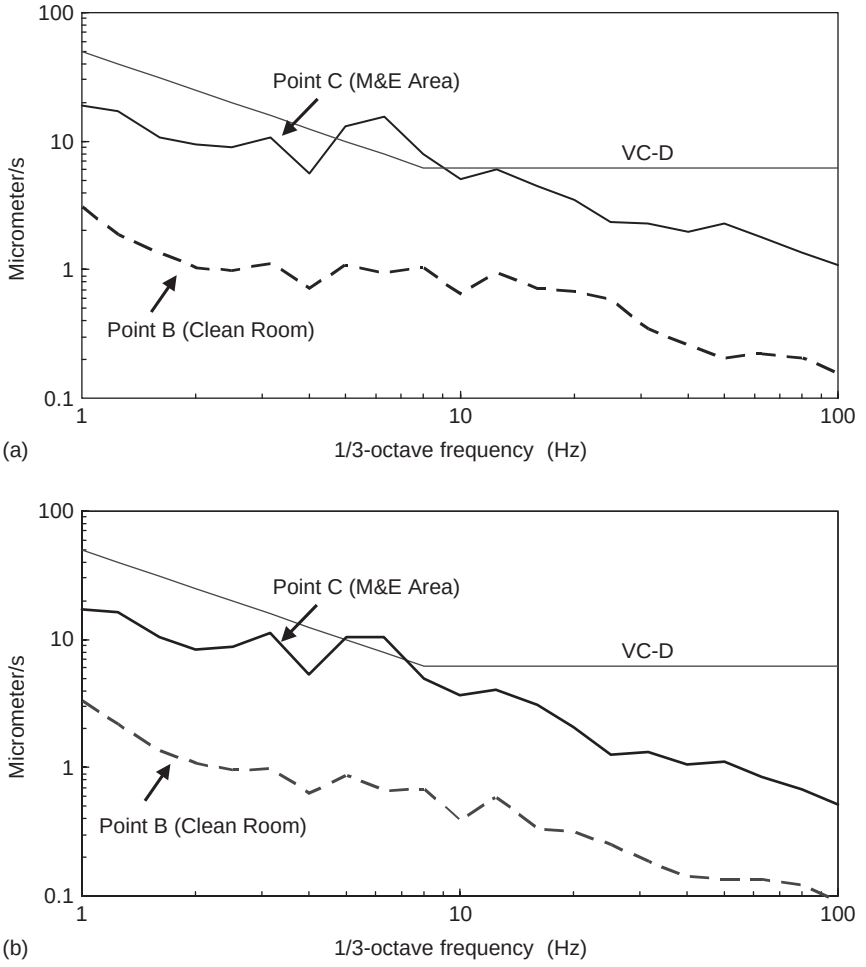


Figure 15 Reduction of vibration response in an expanded area for clean room of an existing building.

are supported on columns and together with the deep two-way ribs provide a very stiff floor system while minimizing dead weight (Amick et al 1991). The holes between the ribs allow air to flow across individual floor levels that it divides to achieve the stringent air cleanliness requirement. Alternatively, a thick flat slab with circular holes can be adopted. When free passage of air between floors is not required, flat slabs without holes supported on a grid of deep beams can be used. The primary design criterion for the floor slab is the vertical vibration induced by human footfall vibration. The floor system should have sufficiently high stiffness and natural frequency to mitigate the adverse effect of human footfall.

For analysis, the finite element model for the waffle slab comprises 3×3 panels of a corner of the production area as shown in Figure 17. The outside edges are unrestrained

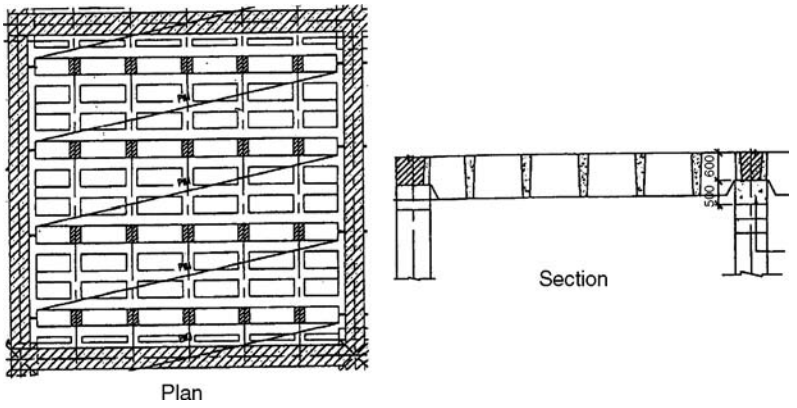


Figure 16 Plan and section of part of a typical waffle floor system.

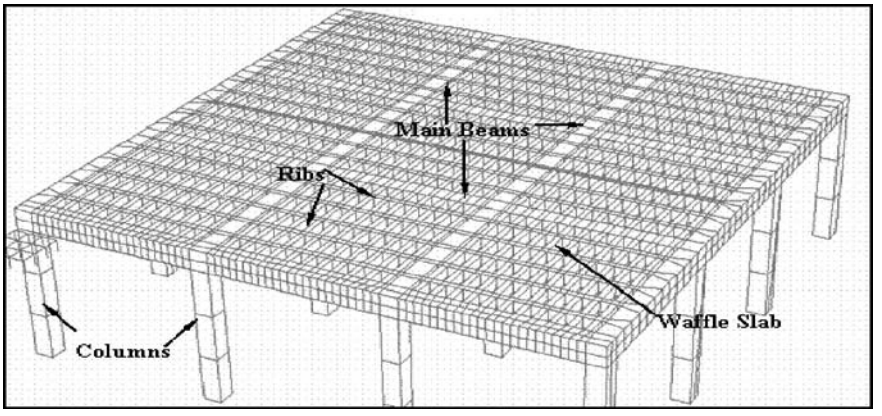


Figure 17 Finite element model of waffle floor system (3 × 3 panels).

except at the column supports while the inner edges, assumed to be lines of symmetry for simplicity of analysis, are restrained in rotation and laterally. The waffle slab is modelled as a grid of beam elements. The raised floor is modelled as an assemblage of thin shell elements with hollow section rod elements for the pedestals. The footfall loading is applied on these shell elements.

Human footfall loading can be modelled as a periodic loading whose parameters within each period depends on the pace and kind of motion. Various representations have been utilized by different researchers and Figure 18 shows several typical footfall approximations obtained by Galbraith and Barton (1970). The period of the pulse varies from about 0.3 sec to 1.5 sec depending on the speed of walking. The general shape is little affected by the weight of the walking person, however, its peak value and rise time depend significantly on walking speed and walker's weight. The periodic repetition may be decomposed into a series of harmonic loads through a Fourier transformation. The third harmonics in the loading spectrum is sufficiently significant and

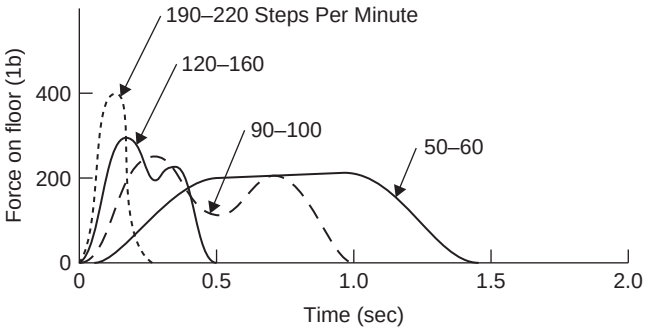


Figure 18 Approximate footfall-induced forces produced by a 75-kg male walker (Galbraith and Barton 1970).

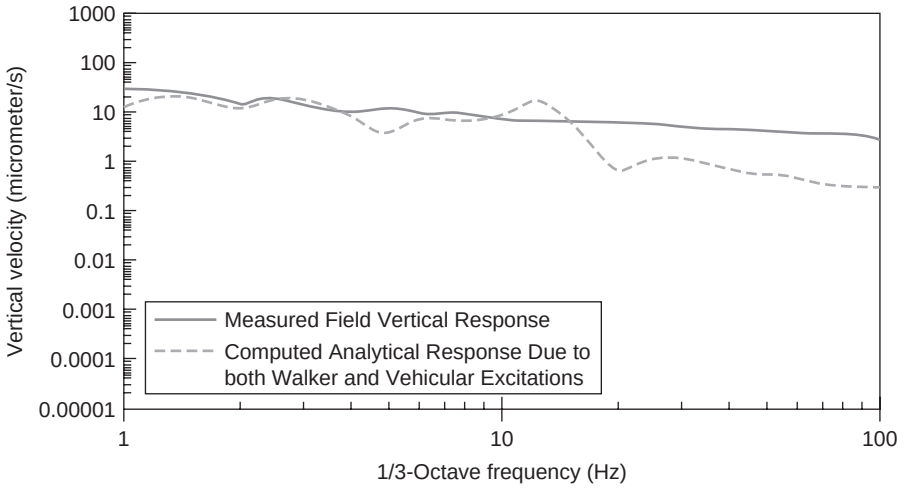


Figure 19 Combined analytical response due to both walker and vehicular excitations.

should be included in the analysis (Bachmann and Ammann 1987; Dowding 1996; Allen 1990; Ellis and Ji 1997).

Vibration measurements on the waffle slab were taken both with and without human walking in a verification measurement. Footfall analysis was performed to simulate 5 persons weighing 70–80 kg walking slowly (about 90–100 steps per minute) around the measurement location to obtain the analytical root mean square (rms) velocity, $V_{rms}(foot)$. The footfall component is combined with the ground borne vibration component due to vehicular traffic excitation to obtain the resultant velocity response $V_{rms}(res)$ as follows:

$$V_{rms}^2(res) = \sqrt{V_{rms}^2(foot) + V_{rms}^2(traf)} \tag{3}$$

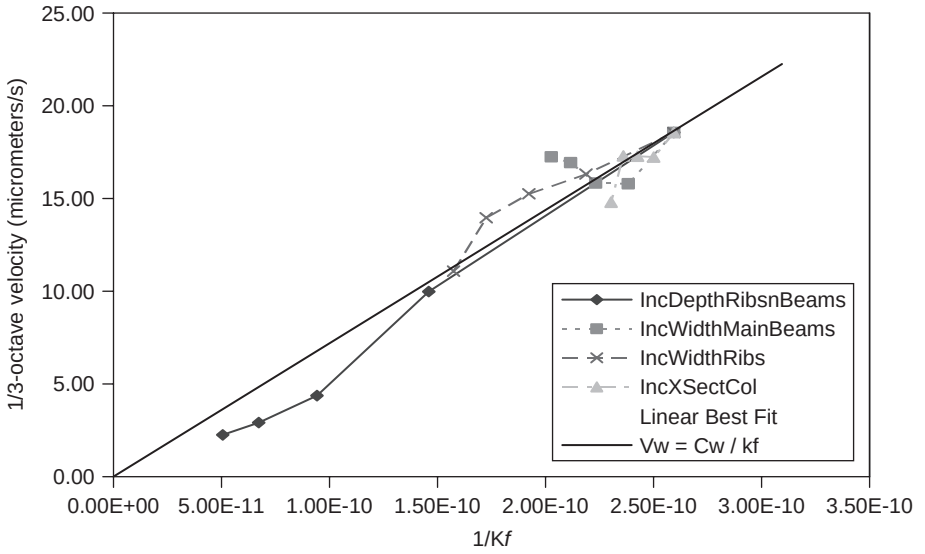


Figure 20 Relationship of 1/3-octave velocity with $1/(Kf)$ (Unit for Kf: is N-Hz/m).

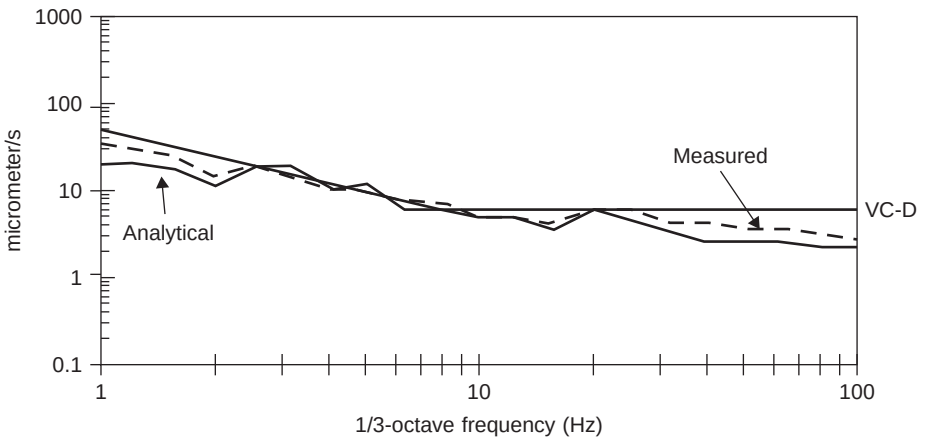


Figure 21 Vertical velocity response in a VC-D area due to footfall.

The resultant response is compared with site measurement in Figure 19. The analytical response obtained for the lower range of frequencies between 1 to 20 Hz agree very well with site measurement, showing slight discrepancy only at a few discrete frequencies. At the 4–7 Hz range, the analytical response is also much lower than field measurement because of the intrusive vibration from a poorly damped compressor plant operating at a frequency of about 7 Hz. At frequencies above 20 Hz, the discrepancies between the field measurement and analytical response are largely due to vibrations contributed by machines and equipment operating within the wafer fabrication structure. This has

also been observed in practice where the vertical response generated by machines often exceeds the response generated by human walking in the frequency range exceeding 20 Hz (Amick and Bayat 1998).

Using the model in Figure 17 and a 75-kg walker, various configurations of depth and width of rib and main beams were analyzed. The respective Kf indexes were determined where K is the point stiffness at the middle of a panel of the floor slab and f is its fundamental frequency. Figure 20 shows the relationship of the peak 1/3 Octave Band rms velocity amplitude (V_w) with $1/(Kf)$. A linear relationship between the two can be approximated as

$$V_w = \frac{C_w}{Kf} \tag{4}$$

where $C_w = 70\,kNs^{-2}$ for the waffle floor system under study and based on the "standard" walker weight of 75 kg used above. Instead of the rigorous analysis as presented above, this linear relationship provides a simple and yet effective guide to design the floor system. As an example, to satisfy the VC-D specification for commonly used waffle floor systems, the Kf index should be in the range of $8–10 \times 10^9$ N-Hz/m. Figure 21 shows typical vertical response measured at a waffle floor system designed for the VC-D specification.

5 Conclusions

Several crucial issues pertaining to the practical structural design of vibration sensitive factories are addressed in this paper. The finite element analysis is used as the main numerical tool for dynamic analysis, particularly in the calibration of traffic forces. The validated numerical model is then used to study the effects of structural design on the vibration response of the production building and the floor system. The numerical study is complemented by field measurements. Ground vibration measurement before construction provides the necessary input for ground borne study, and building vibration measurement after construction serves as verification. The innovative isolation-and-shield strategy has been found to perform well for many industrial projects for mitigation of ground borne vibrations, wind induced vibrations, and in some cases where needed, M&E induced vibrations. The Kf-index approach correlates well with the finite element results and offers a quick way of designing the floor system for the vibration sensitive area.

Acknowledgement

The authors would like to thank Mr. E K Lau and Mr K M Lee for their assistance in the finite element analysis.

References

Allen, D.E., Building Vibrations from Human Activities. *Concrete International Design and Construction* 12:6, 66–73, 1990.

Amick, H. and Bayat, A., Dynamics of Stiff Floors for Avanced Technology Facilities. *Proceedings of 12th ASCE Engineering Mechanics Conference*, La Jolla, CA, 318–321, 1998.

Amick, H., Hardash, S., Gillett, P. and Reaveley, R.J., Design of Stiff, Low-Vibration Floor Structures. *Proceedings of International Society for Optical Engineering (SPIE)*, San Jose, CA, Vol. 1619, 180–191, 1991.

Bachmann, H. and Ammann, W., Vibrations in Structures Induced by Man and Machine. IABSE-AIPC-IVBH, Vol. 3e. International Association for Bridge and Structural Engineering (IABSE), Zurich, Switzerland, 1–43, 1987.

Dowding, C.H., *Construction Vibrations*, Prentice Hall International Series in Civil Engieering and Engineering Mechanics, 379–39, 1996.

Ellis, B.R. and Ji, T., Human-structure interaction in vertical vibrations. *Structures and Buildings*, *the Proceedings of Civil Engineers* 122:1, 1–9, 1997.

Galbraith, F.W. and Barton, M.V., Ground Loading from Footsteps. *Journal of Acoustical Society of America* 48, 1288–1292, 1970.

Gordon, C.G. and Ungar, E.E., Generic Vibration Criteria for Vibration-Sensitive Equipment. *Proceedings of International Society for Optical Engineering (SPIE)*, San Jose, CA, 1619, 71–85, 1991.

Koh, C.G., *Modelling of Dynamic Response of Elastomeric Isolation Bearings*, Ph.D. Thesis, University of California, Berkeley, 1986.

Leung, K.W. and Papadimos, C. A., *Micro-Vibration Criteria for 300 mm and Beyond. Semiconductor Fabtech* 10th edition, Henley Publishing, London, 1999.

Northwood T.D., Isolation of Building Structures from Ground Vibration. *Isolation of Mechanical Vibration, Impact and Noise, AMD,* 1, 87–96, 1977.

Ungar, E.E. and White R.W., Footfall Induced Vibration of Floors Supporting Sensitive Equipments. *Sound and Vibration,* 10, 1979.

Chapter 2

Fatigue and durability of steel bridges

J.W. Fisher & S. Roy

ATLSS Research Center, Lehigh University, Bethlehem, PA, USA

ABSTRACT: The historical performance of welded steel bridges in the USA is reviewed with focus on fatigue cracking at welded steel details, as well as distortion induced cracking at web gaps. Also examined is the impact of corrosion loss of section as a contributing time-dependant factor to fatigue resistance. The role of steel materials including modern high performance steels (HPS) is reviewed as is the influence of corrosion. Variable amplitude loading is also examined and compared with current design of steel bridges for fatigue resistance. Also examined are methods to improve and retrofit fatigue sensitive details, including modern post-weld enhancement by ultrasonic impact treatment (UIT).

1 Introduction

Welded and bolted details for bridges are designed based on the nominal stress range rather than the local "concentrated" stress (hot spot) at the weld detail. The nominal stress is calculated with mechanics of material equations and does not include the local effect of stress concentrations of welds and attachments. Since fatigue is typically only a serviceability problem, members are designed for fatigue using service loads that represent the variable load spectrum. It is standard practice in fatigue design of welded structures to separate the weld details into categories having similar fatigue resistance in terms of the nominal stress. Each category of weld details has an associated S-N curve. The S-N curves for steel highway and railway bridge details (AASHTO 2004, AREMA 2005) are shown in Figure 1. The S-N curves are presented for seven categories of weld details A through E′, in order of decreasing fatigue strength. These S-N curves are based on a lower bound to a large number of full-scale fatigue test data with a 97.5% survival limit, as illustrated in Figure 2, which shows the fatigue test data for Category E and includes a wide variety of cover plate details (Fisher et al. 1970). It should be noted that the variability is related to the cycle life, which is described by a log normal distribution at all stress range levels for a given detail.

Figure 1 shows the fatigue thresholds or constant amplitude fatigue limits (CAFL) for each category as horizontal dashed lines. When constant amplitude tests are performed at stress ranges below the CAFL, detectable cracking does not occur. The CAFL occurs at an increasing number of cycles for lower fatigue categories or classes. Sometimes, different details, which share a common S-N curve (or category) in the finite-life regime, have different CAFL.

Typically, small-scale specimen tests will result in longer apparent fatigue lives particularly near the CAFL. Therefore, the S-N curves were based on tests of full size

structural components such as girders (Keating & Fisher 1986). Testing on full-scale welded members has indicated that the primary effect of constant amplitude loading can be accounted for in the live-load stress range, i.e., the mean stress is not significant. The reason that the dead load has little effect on the lower bound of the results is that, locally, there are very high residual stresses from welding. Mean stress may be important for some details that are not welded, however. In order to be conservative for non-welded details, in which there may be a significant effect of the mean stress, the fatigue test data should be generated under loading with a high tensile mean stress.

Cover plate details and equivalent details such as gusset plates are common on older bridges. Hence it is necessary to examine the service loading that bridges are subjected to.

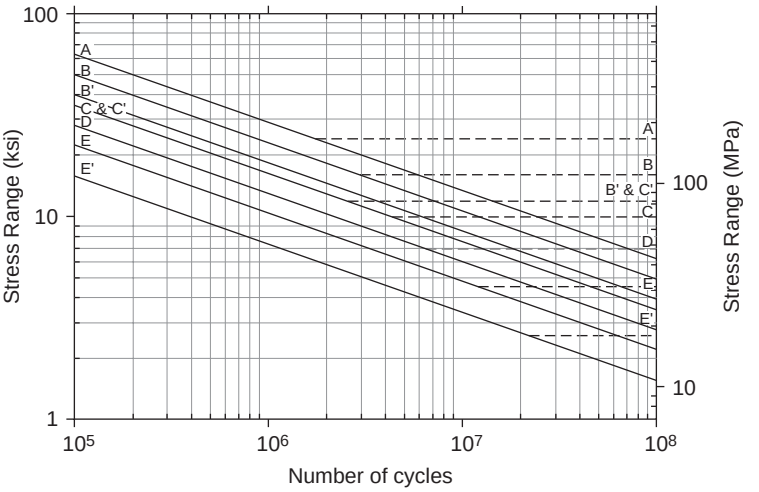


Figure 1 AASHTO design S-N curves in use since 1983.

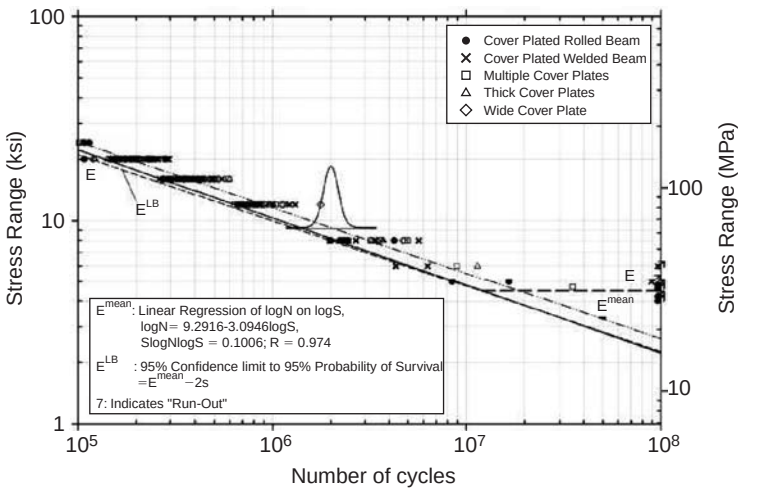


Figure 2 97.5% survival S-N curve for Category E details.

2 Variable-amplitude fatigue

All service load histories for bridges consist of stress range cycles of varying amplitudes, hereafter called variable-amplitude (VA) loading. However, the design curves of stress range versus number of cycles, commonly called design S-N curves in the AASHTO LRFD bridge design specifications (2004), are based on fatigue tests that were performed under constant-amplitude (CA) loading. Since stress ranges vary in size and numbers, the relationship between CA and VA loading should be understood.

A histogram of VA stress ranges can be converted into an equal number of CA stress ranges that produce the same amount of crack growth and damage as is the case for the VA stress ranges. This so called equivalent CA stress range is based on Miners Rule and given by:

$$S_{re} = \left[\sum_i \varphi_i \cdot S_{ri}^3 \right]^{1/3} \tag{1}$$

where, S_{ri} = the *ith* stress range in the histogram; ϕ_i = frequency of occurrence of S_{ri}; and 3 is the slope of the log-log linear S-N line for CA fatigue. This equivalent stress range of S_{re} is referred to as the root-mean-cube (RMC) stress range. Schilling, Klippstein, Barsom, & Blake (1978) investigated both the RMC and the root-mean-square (RMS) methods of assessing the variable stress range spectrum. In those tests the stress cycles exceeding the CAFL had frequencies that were 10% or greater of the cumulative frequency.

Figure 3 shows a schematic of an S-N curve with variable stress-range distributions added at three different levels. When stress ranges in the variable distribution are higher than the CAFL, and have a frequency of 0.05% or more of the cumulative frequency,

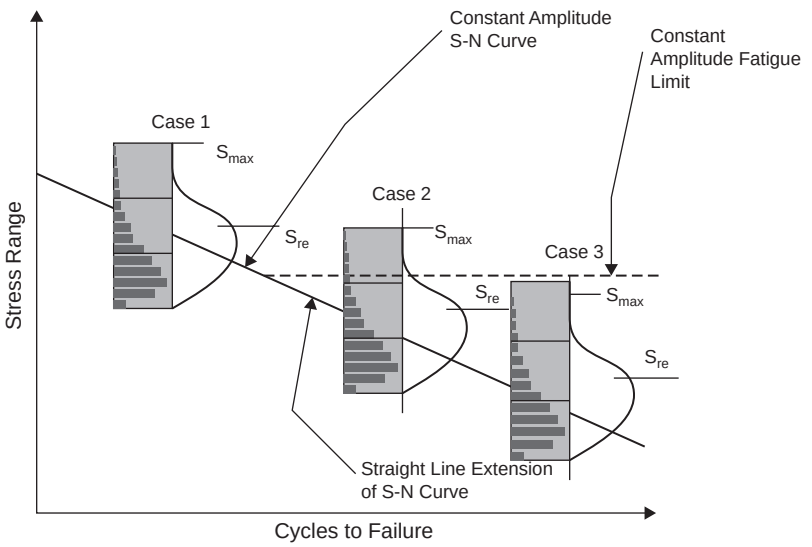


Figure 3 Application of design S-N curves under variable amplitude stress spectrum.

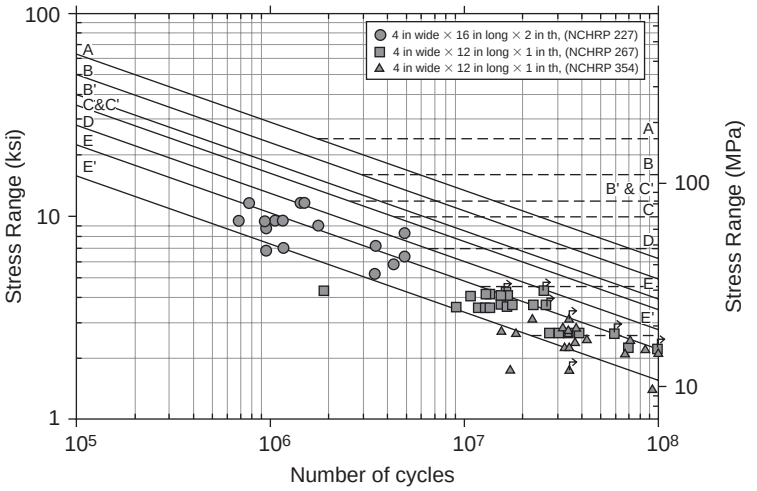


Figure 4 Fatigue performance of web attachments under VA loading.

the equivalent stress range S_{re} is used with the log-log linear S-N curves the same way as a CA stress range would be used.

Tilly & Nunn (1980) tested 14 tensile specimens with category E longitudinal gusset plates under CA cycling and seven specimens under VA cycling with a Rayleigh spectrum. The tests were performed in a resonant fatigue machine at a frequency of 150 Hz. With fatigue lives of up to 415 million cycles, these tests are the longest on record and the only ones carried out into the infinite-life regime of category D, E and E′ details. CA and VA stress range data correlated well, and the results confirmed the existence of infinite life if all stress ranges are below the CAFL.

Fisher et al. (1983) tested eight beams under VA loading, also with a Rayleigh spectrum. Although the tests were carried out to 150 million cycles, the data fell in a wide band along the sloped S-N line of the finite-life regime. Following up on this study Fisher et al. (1993) tested eight more beams. Each beam had two Category E′ cover plates, six Category E′ web gusset plates and three Category C′ transverse stiffeners. As in the earlier study, the VA loading corresponded to a Rayleigh spectrum. All data points for the category E′ cover plates and gusset plates fell in the finite life regime. Two of 24 stiffeners, with fatigue lives of 109 and 120 million cycles, provided some information on the variable amplitude fatigue limits (VAFL). The Category E′ test data including all stress cycles greater than 10 MPa from these two studies are shown in Figure 4 and compared with the fatigue design curves. It is apparent that the test data all follow the straight line extension of the resistance curve below the CAFL.

These studies demonstrated that under long-life variable amplitude loading the cumulative frequencies of cycles which exceeded the CAFL by more than 0.05% resulted in fatigue cracking of welded web attachments. Damage accumulation was observed to occur from all stress cycles that exceeded 10 to 15 MPa. The tests on stiffeners which were Category C′ details, only developed cracks in two out of 20 tested

when 0.01% of the stress cycles exceeded the CAFL. Altogether up to 120 million variable cycles were accumulated in these stiffener details.

3 Fatigue design truck

The fatigue design truck in the LRFD specifications is based on the design truck, which consists of 35 kN front axle and two 145 kN rear axles. The axle spacings are 4.3 m on the tractor and 9 m on the trailer. Two assumptions were made: (1) the stress range induced at a detail by a truck crossing the bridge is proportional to the gross truck weight; and (2) all trucks crossing the bridge produce the same fatigue damage as is done by an equal number of trucks with axle loads of 75% of the design truck, and axle spacings of 4.3 and 9 m. Analogous to calculating the equivalent stress range of a VA spectrum with Equation 1, the equivalent fatigue truck weight was calculated as (Moses et al. 1987):

$$W_e = \left[\sum_i \varphi_i \cdot W_i^3 \right]^{1/3} \tag{2}$$

where W_i = weight of the *ith* truck in the histogram; ϕ_i = frequency of occurrence of W_i. A slope of 3 for the log-log linear S-N line was used to establish the effective weight. Equation 2 yields the so called RMC truck weight. The distribution used to calculate the fatigue truck weight comes from a study by Snyder et al. (1985) in which a total of 27,513 trucks were weighed on 31 Interstate, US route and State route bridges located in seven states. Truck weights were distributed as follows: 99.47% in 25–445 kN, 0.38% in 445–667 kN, and 0.05% in 667–890 kN. Applying Equation 2 to the weigh-in-motion data of Snyder et al. gives an equivalent truck weight $W_e = 239$ kN or, relative to the HS20 design truck weight: $239/325 \cong 0.75$. This value is not a load factor in the sense of LRFD as represented in Table 3.4.1-1 and Article 6.6.1.2.2 of the LRFD bridge design specifications (2004). In Article C6.6.1.2.2, the load factor for fatigue is acknowledged to be 1.0. The absence of a load factor is related to field experience, which shows that measured stress ranges are always smaller than calculated stress ranges.

A constant axle spacing of 9 m was found to best approximate the axle spacing of typical four and five axle trucks responsible for most fatigue damage to bridges. Clearly, most structures carry enough truck traffic to justify designing them for an infinite fatigue life, especially the deck elements. In most cases, designing for infinite fatigue life adds little to cost.

4 Experience with fatigue cracking

Fatigue cracking in steel bridges in the United States has become more frequent in its occurrence since the 1970's. A large crack was discovered in 1970 at the end of a cover plate in one of the Yellow Mill Pond multi-beam structures located at Bridgeport, Connecticut (Fisher 1984). Between 1970 and 1981, numerous fatigue cracks were discovered at the ends of cover plates in this bridge as seen in Figure 5. In most cases, fatigue cracking in bridges resulted from an inadequate experimental base and overly optimistic specification provision developed from the experimental data in the 1960's.

Figure 5 Yellow Mill Pond Bridge: fatigue crack at the end of cover plate.

Subsequent laboratory data has verified the low fatigue strength in the high cycle region. The assumption of a fatigue limit at 2×10^6 cycles proved to be incorrect. As a result of extensive large-scale fatigue testing, it is now possible to clearly identify and avoid details which are expected to have low fatigue strength. The fatigue problems with the older bridges can be avoided in new construction if good detailing practice is followed and each detail is designed such that the stress range due to applied live load is below the design allowable stress range. Fortunately, it is also possible to retrofit or upgrade the fatigue strength of existing steel bridges with poor details. Fatigue crack repair techniques for weld toe surface cracks were developed to repair and extend the life of cracked welded steel bridge details.

Peening works primarily by inducing a state of compressive residual stress near the weld toe (Fisher et al. 1979). Because the benefit of peening is derived from lowering the effective tensile stress range, it has been found to be the most effective when conducted under dead load. In this case, the peening only needs to be effective against live load. Air-hammer peening (AHP) can be a successful repair as long as the crack depth does not exceed the zone of compressive stress. The depth of compressive stress is maximized by using air pressure lower than 290 kPa and up to six passes with a peening tool. Fatigue cracks up to 3 mm deep and 50 mm long at the cover plate weld toe can be arrested by peening provided the stress ranges do not exceed 40 MPa. Peened beams with crack depths larger than 3 mm usually show no measurable increase in life and other repair procedures such as bolted splices may be required.

Figure 6 shows a typical peened weld toe from the Yellow Mill Pond Bridge. The severity of the deformation is indicated by the elongation of the grains, which decreased with the depth below the surface. Fatigue tests of these full size cover plated beams are plotted in the S-N relationships provided in Figure 7. This structure was retrofitted by peening in the 1970's. This prevented subsequent crack growth in this heavily used structure until it was replaced in 1997. Subsequently, several beams removed from the original structure were tested in the laboratory (Takamori & Fisher 2000). These tests were carried out at a stress range of 70 MPa which exceeded the maximum stress range in the variable amplitude spectrum that the bridge was subjected to for over 20 years

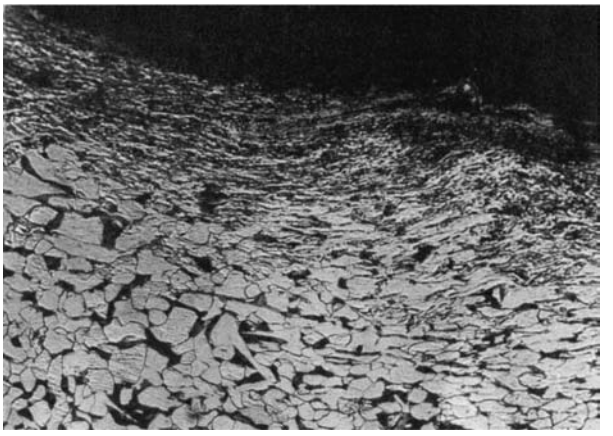


Figure 6 Air hammer peened cover plate end weld toe.

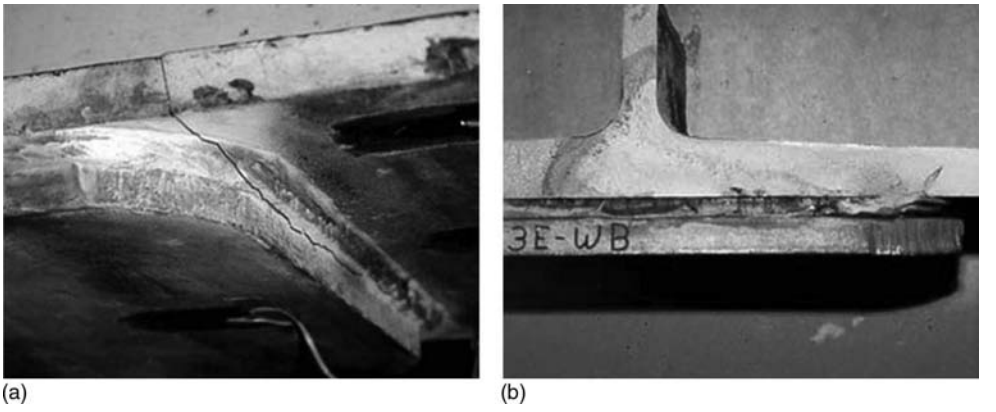


Figure 7 Through throat cracking at peened cover-plate end weld in Yellow Mill Pond Bridge: (a) fatigue crack progression; (b) exposed fracture surface.

of service after treatment. The tests verified that no fatigue crack growth had occurred in these bridge details after more than 60 million cycles of truck loading. The peened and gas tungsten are re-melted retrofitted details had successfully prevented further growth at the treated details. At the higher test stress range, the laboratory tests were found to develop fatigue cracks in the weld throat as illustrated in Figure 8 and not at the treated weld toe. Figure 9b shows another successful retrofit by peening a fillet weld joining a gusset plate to a floor beam flange. The small crack seen in Figure 9a was ground out prior to peening.

5 Distortion-inducted fatigue

Most cracks found in bridges were caused by distortion of member cross sections and out-of-plane deformations of webs that induced localized bending stresses (Fisher et al.

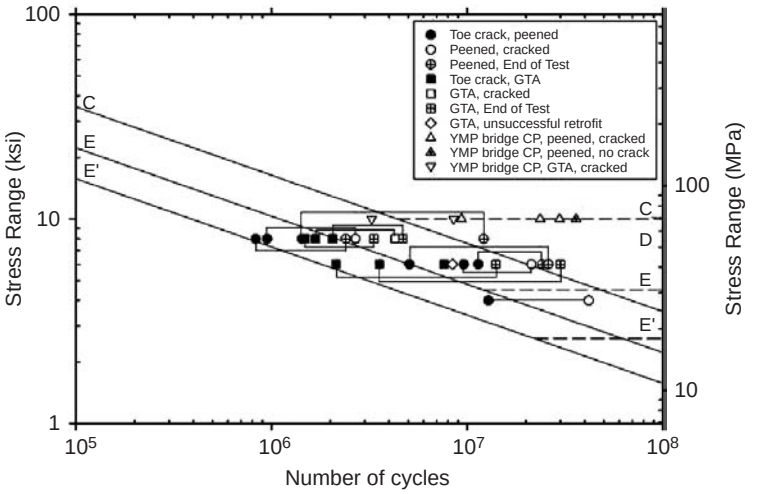


Figure 8 Results of retrofitted full size cover plated beams.

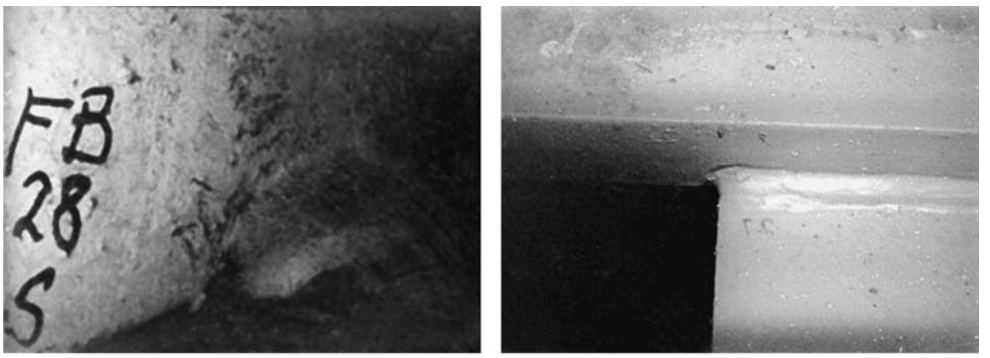


Figure 9 Retrofitting by peening: (a) small fatigue crack at weld toe of gusset plate; (b) crack suppressed by grinding followed by peening.

1990). Out-of-plane distortion occurs mainly at attachment plates for diaphragms, transverse stiffeners and floor beams as illustrated in Figure 10. The solution to this problem lies in proper detailing that eliminates the secondary stresses.

In older bridges, transverse stiffeners and attachment plates were not welded to the tension flange of welded I-girders and box girders for fear that a fatigue crack initiating in the flange would lead to a brittle fracture. This well-intended but outdated practice originated in Europe in the 1930's from unexpected brittle fractures in early welded bridges, which was attributed to the welded details, but was primarily due to the poor quality of the steel. This practice has been the main cause of distortion-induced cracking, which can be prevented by welding stiffeners to the web and the flanges. Figure 11 shows a crack that formed along the fillet weld that attaches a diaphragm

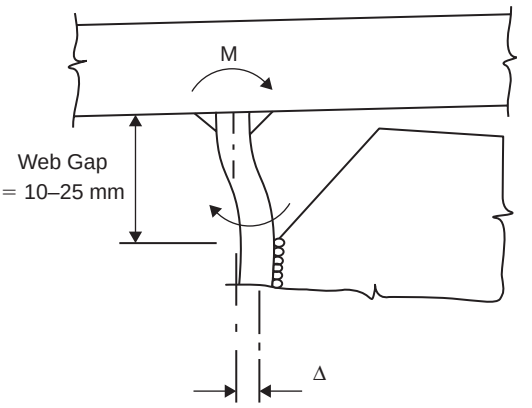


Figure 10 Schematic of web gap distortion.

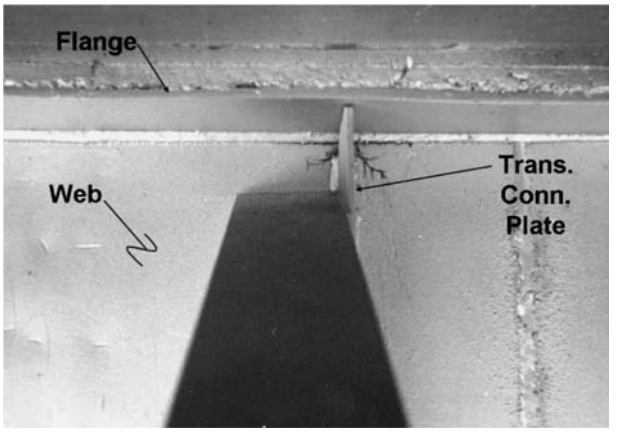


Figure 11 Distortion induced fatigue cracking in the web gap.

connection plate to the web of a plate girder. The fatigue cracking in these gaps typically occurs in a longitudinal direction along the fillet weld toe of the longitudinal web-to-flange joint, at the termination of the vertical fillet weld, or at both locations, as shown in Figure 11. In many cases, the displacement causing this distortion is limited and the cracking arrests as the compliance of the gap increases and the stresses are reduced. Because most of the cracks are oriented longitudinally, there is typically no reason to be concerned about fracture of the girder unless the cracks turn downward and propagate across the web. It is typically a mistake to try to weld repair such cracks, as this restores the high stresses which originally caused the cracking and will certainly reinitiate the cracking at the weld repair.

In the case where the distortion is limited, holes may be drilled or cored at the crack tips to temporarily arrest propagation. Figure 12 shows holes used to temporarily arrest crack growth. In this case the cracks reinitiated because the out-of-plane bending

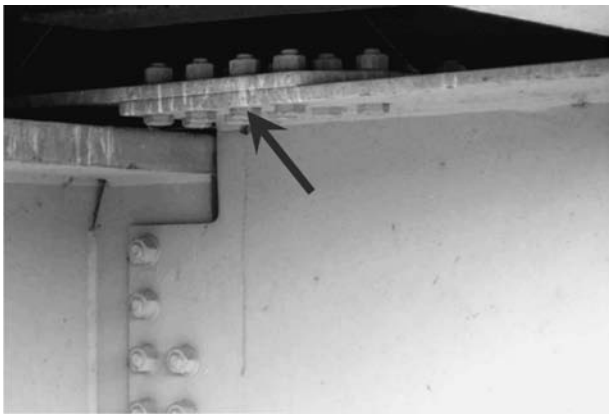


Figure 12 Retrofitting connection by drilling holes.

remained excessive. Holes essentially blunt the tip of the crack. Experimental studies have indicated that the size of the hole must satisfy the relationship.

$$\frac{\Delta K}{\sqrt{\rho}} = \frac{S_r\sqrt{\pi a}}{\sqrt{\rho}} < 10.5 \cdot \sqrt{\sigma_y} \tag{3}$$

where ρ = radius of the hole (mm); $2a$ = assumed crack size (mm), defined between the edges of the hole (Fisher et al. 1980); K = stress intensity factor for a crack size of $2a$; and σ_y = yield strength of the plate (MPa). The validity of this equation has been studied in the laboratory on full-scale welded beams subjected to variable amplitude loading up to 90 million stress cycles.

In most cases, problems with web-gap-cracking can be solved by rigidly connecting the attachment plate to the tension flange. To retrofit existing bridges, a rigid tee or angle may be connected to the attachment plate and the tension flange using high-strength bolts. Holes must also be drilled at the ends of short cracks as a temporary means of extending the fatigue life. The cracked detail shown in Figure 12 was retrofitted in this way.

In cases where the distortion is displacement limited, the stresses can be reduced by increasing the flexibility of the connection. The flexibility may be increased by allowing a small fatigue crack to remain after drilling or coring holes at the crack tips. Another way to increase the flexibility of the joint is to remove part of the stiffener or connection plate and increase the length of the gap, as shown in Figure 13. If the same displacement occurs over a greater length gap, the stresses are significantly reduced. Note also the holes drilled at the crack tips along the web-flange weld.

In case where the distortion is not displacement limited (i.e. load controlled), hole drilling or increasing the flexibility of a connection will not work. In these cases, and even in many displacement limited cases, the best solution to distortion cracking is to increase the rigidity of the connection. In new construction, the bridge specifications now recommend that stiffeners and connection plates should be rigidly connected to

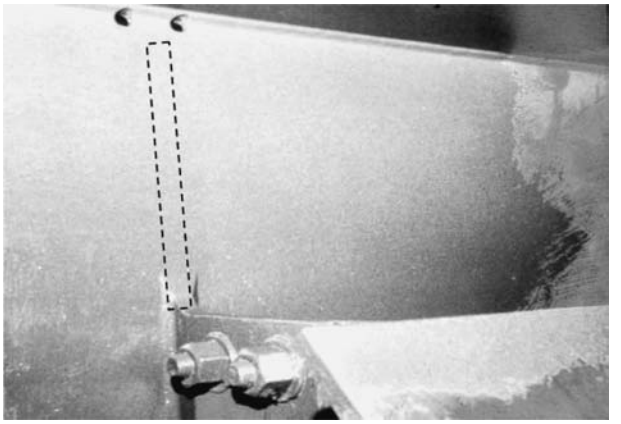


Figure 13 Retrofitting by connection softening.

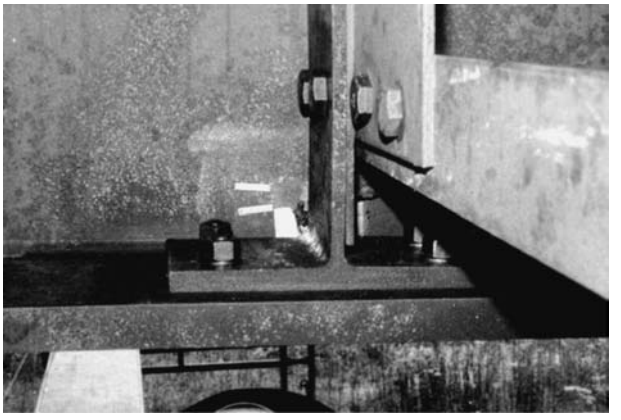


Figure 14 Retrofitting by connection stiffening.

both flanges and the web. Figure 14 shows a tee connection bolted to the flange and transverse connection plate in order to "bridge" the web gap.

The December 2000 fracture of the Hoan Bridge in Milwaukee, WI, seen in Figure 15, also involved a lateral gusset plate that was welded to the web but not to the transverse stiffeners. The details for the center and outside girders are shown in Figure 15. The gusset plate was fitted tightly against the transverse stiffener creating a crack-like geometric condition at the intersection of the longitudinal welds attaching the gusset to the web and the transverse stiffener welds (Figure 16). This resulted in a zone of high triaxiality in the girder web. In addition, the indirect bolted connections between the gusset plate and transverse stiffener were flexible. Cleavage fracture developed in the girder web without detectable fatigue cracking as a result of the triaxiality and crack-like geometry as seen in Figure 17. The detail was not inspectable, and examination could not reveal any crack extension prior to fracture

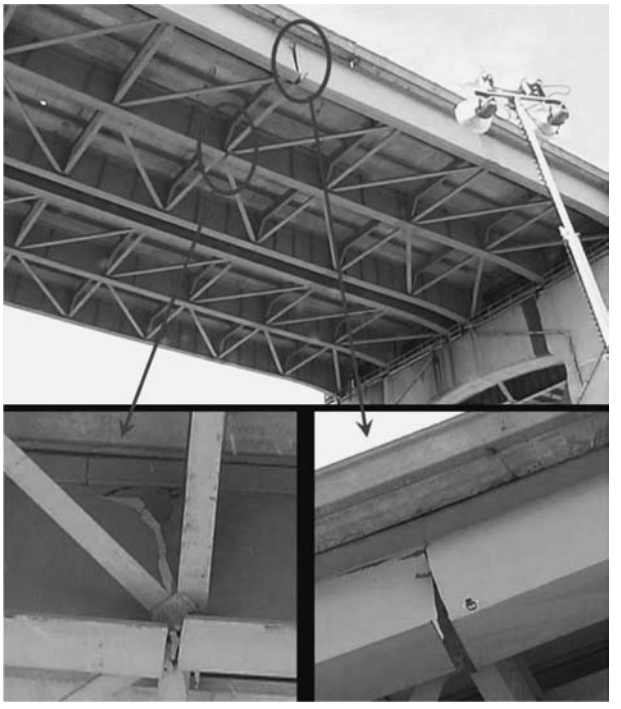


Figure 15 Hoan Bridge fracture.

(Wright et al. 2003). 100 mm holes were drilled at all details to isolate the crack initiation zone. This permitted limited traffic to resume using part of the structure pending final repair and retrofit which involved removing all of the bottom laterals and gusset plates (Sivakumar et al. 2003).

The development of these cracks indicated that considerable care must be exercised when web gusset plates are used for the bottom lateral system. No intersecting welds should be permitted, and ample copes should be utilized to prevent small web gaps. Positive connections must be provided between horizontal gusset plates and the transverse connection plates. It seems preferable to bolt lateral gusset plates to the bottom flanges, or alternatively avoid the use of bottom lateral systems.

6 Effect of corrosion on fatigue life

6.1 *Corrosion notching*

Severe notches can develop in sections, particularly where dirt and debris accumulate and create an active corrosion cell. This type of notching is not restricted to weathering steel but will occur in all steel materials. The fatigue resistance of corrosion-notched riveted sections has been examined in full-scale members removed from bridges Fisher et al. 1987, Fisher, Yen & Wang 1990. Figure 18 shows corrosion notched flange angles of riveted members where dirt and debris accumulated under a diaphragm and

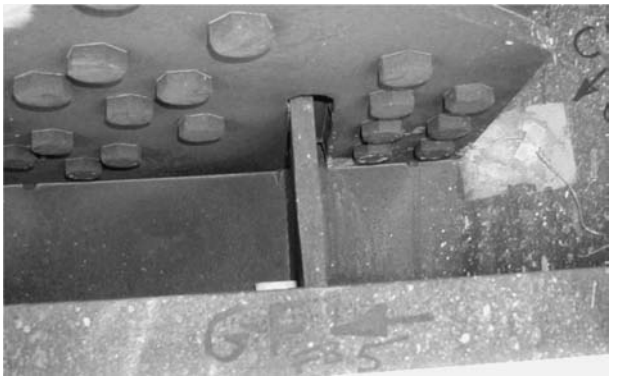


Figure 16 Constraint at the intersection of gusset, transverse stiffener and girder web.

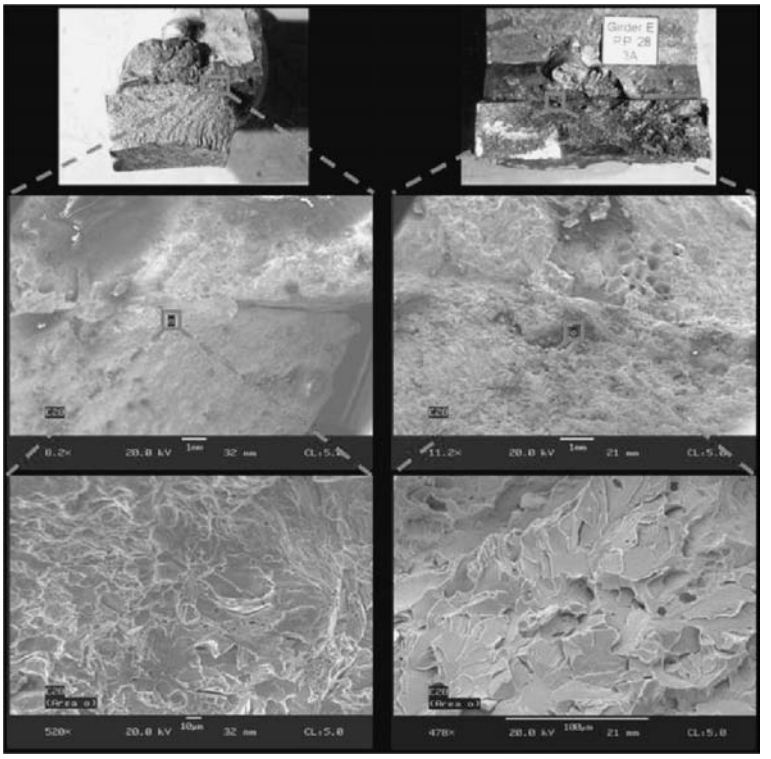


Figure 17 Triaxiality induced cleavage fracture.

provided an active corrosion cell site. Figure 19 shows a crack that developed in the corroded web plate of the rolled section.

Corrosion notches provide a substantial reduction in fatigue resistance in laboratory tests. As was demonstrated by Fisher et al. (1987) and Fisher, Yen & Wang (1996),

cracks developed in the gross section at the notches with little or no influence of rivet holes when more than 50% of the thickness was lost. The fatigue resistance of the severely corrosion-notched section was observed to be as low as Category E.

It is interesting to note that in actual structures cracks are seldom observed at the corrosion-notched section until significant loss in the total section occurs, apparently because of the fact that the corrosion process is continuous and does not permit a crack to develop. Any local damage from cyclic loading appears to be removed by the ongoing corrosion process. Cracks form only as the corrosion loss becomes extensive.

6.2 *Corrosion effects on boundary conditions*

An example is the significant role that crevice corrosion played n the development of the collapse of the Mianus River Bridge carrying I-95 in 1983. (Demers & Fisher 1990,

Figure 18 Corrosion notch in flange angle of riveted member under a diaphragm.

Figure 19 Fatigue crack in heavily corroded flange angle and web plate.

NTSB 1983) Figure 20 shows the pin and remaining hanger that had supported the suspended span. A crack undetected before the collapse resulted in the fracture of a segment of the pin supporting the hanger, as illustrated in Figure 21. Final collapse immediately followed. Corrosion and the resulting geometric changes at the pin connection played a major role in the development of the crack in the pin. A cross section through the center pin, washers, and girder web is shown in Figure 22. The corrosion packout between the spacer washers and the hanger forced the outside hanger toward the end of the pin, causing yielding and crack initiation in the pin and hangers (Fisher & Pense 1984).

Although close visual inspection would likely provide an indication of geometric change and the existence of corrosion packout, other inspection tools and sensors are

Figure 20 Collapse of suspended span of Mianus River Bridge, Connecticut.

Figure 21 Pin fracture segment.

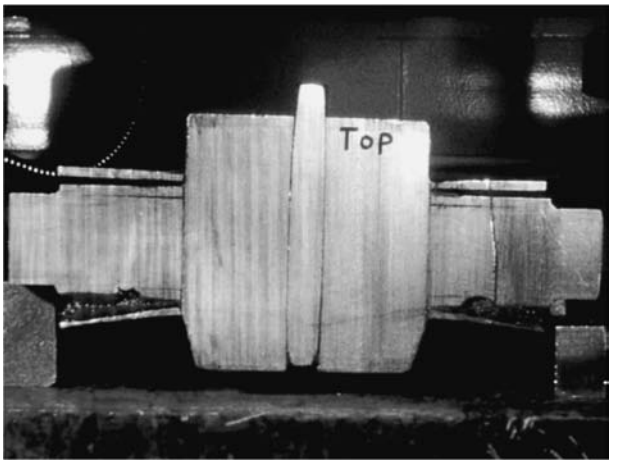


Figure 22 Cross-section through pin, web and spacer washers showing corrosion packout.

needed to identify active corrosion cells as well as the crack initiation and the observed unusual displacements.

Another example of corrosion-related cracking is the fixity that occurs at pin connections. Debris and contaminated water cause corrosion, which can lead to increased joint fixity and result in cyclic stresses that are up to an order of magnitude greater than the design stress range. Figure 23 shows the hanger of a deck truss that developed fatigue cracks in the riveted built-up sections (Demers & Fisher 1990). Strain measurements demonstrated that the pin fixity caused bending stress as well as sudden joint release, resulting in stress range conditions well above the fatigue resistance of riveted joints. Figure 24 shows a close up of the fatigue crack that initiated from a rivet hole in the web plate. The crack was arrested with a drilled hole, also shown in the picture (to the right).

6.3 *Weathering steel*

Atmospheric corrosion can lead to deterioration of steel structures, depending on the environment. Weathering steel has been developed as a low-alloy material that promotes the formation of a protective oxide coating (rust layer) as a self-healing barrier against further corrosion (Shastry et al. 1987). After a few years, the thickness loss of A588 and A242 weathering steels shows a decrease in the corrosion rate. After 16 years of exposure, the thickness loss 90 μm (<4 mils). Hence, the corrosion loss is negligible provided dirt and debris are prevented from accumulating and forming a poultice and corrosion cell.

The fatigue characteristics of welded details in weathering steel were examined by Albrecht & Naeemi (1984), Albrecht et al (1989) and Yamada & Kikuchi (1984). Both test series used plate specimens fabricated with automatic submerged arc welds, which resulted in very good profiles. The specimens ere tested under constant-amplitude

Figure 23 Riveted deck truss hanger with fatigue cracking from pin fixity.

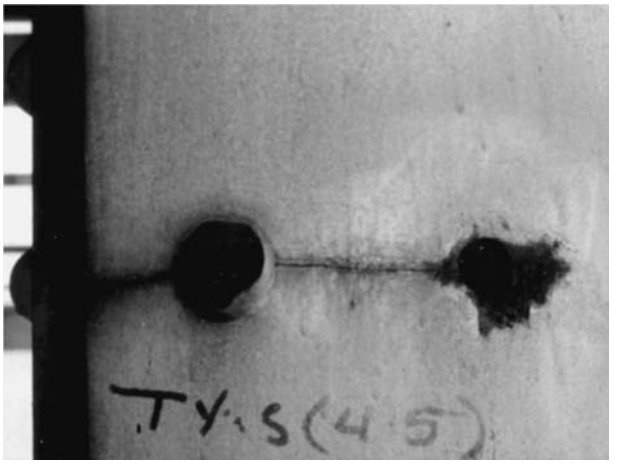


Figure 24 Close-up of fatigue crack from rivet hole with crack arrest hole.

load conditions. Each program examined non-weathered and weathered specimens. The weathered specimens were subjected to varying degrees of atmospheric exposure before testing. The tests did not simulate actual field conditions because the weathering process was not continued during the actual fatigue testing.

The study by Albrecht and his colleagues involved the fatigue testing of specimens that simulated either a transverse stiffener detail or an attachment plate. The simulated stiffener specimens were 25 mm and 10 mm thick plates, smaller than similar cruciform specimens used in other studies in the as-received condition and after 2 and 4 years of exposure or 3 and 8 years of exposure. All test results exceeded the Category C

resistance curve. Other test results at and below the Category C curve were from non-weathered steel specimens reported by Kunihiro et al. (1972) and Keating and Fisher (1986). The attachment specimens consisted of a 100 mm long plate welded around the entire perimeter to a base plate. This detail would normally correspond to Category D. The specimens were fatigue tested as fabricated (non-weathered), after 2 years of exposure, and after they had been weathered for 4 years. The results from this study also plotted significantly above the Category D resistance curve, just as were observed with other small-scale simulated tests.

The study by Yamada & Kikuchi (1984) used two different types of plate specimens: non-load-carrying cruciform joints and gusset-plates. Both weathering steel and standard structural steel were used for the specimens. The cruciform specimens consisted of two transverse attachments welded to a 13 mm × 75 mm base plate. The gusset specimens were fabricated with two longitudinal attachment plates, each 100 mm long, welded along the edge to the base plate. The specimens were fatigue tested as fabricated, after they had weathered for 2 and 4 years of exposure. In addition, stiffener details were cut out of the web of an actually weathered steel bridge that had been in service for approximately 5½ years. All cruciform test results plotted beyond the Category C curve. The results of the gusset plate specimen also fell above the Category D resistance curve.

A full-scale A588 steel W36 × 230 beam with cover plates 32 mm × 305 mm was subjected to 56 million stress cycles at a stress range of 218 MPa with no detectable cracking (Fisher et al 1979). This specimen was subsequently stored outside the laboratory for a 2 year period. To simulate exposure to deicing salts, road salt was added to a bucket of water and the suspended solution was poured over the cover plate ends at monthly intervals. Undissolved particles were left in place, and the cover plate ends were exposed to the natural weathering process. Generally, the undissolved particles remained in place for only a few days as they were dissolved and washed away in rain. The girder was subjected to a cyclic stress range of 41 MPa. Cracking through the weld throat was noted after 7.1 million cycles at one end and 7.6 million cycles at the second end. Testing was stopped, and the crack was exposed to reveal its shape. No crack tip had propagated into the girder flange. The residual life for the crack to move through the longitudinal welds and through the flange, as was observed with other cover-plated beams that experienced similar throat cracks, was estimated to be an additional 4.4 million cycles. The test results are plotted in Figure 25 and compared with the results of other full-scale cover plated beams. The weathered details exhibited higher fatigue lives than any of the non-weathered specimens tested at the 41 MPa stress range and experienced fatigue cracking. Examination of the cover plate end weld toe in the weathered test specimen showed a decreased stress concentration similar to that observed by Yamada and Kikuchi (1984).

The test data on weathering steel indicate that no significant differences exist between the lower-bound fatigue resistance of weathered and non-weathered steel details. None of the test data for the weathered steel specimens plot below the lower-bound design category applicable to the respective welded detail. Rusting and pitting, as would be expected, introduce notches and can affect the crack initiation behavior of the base metal. Studies on weathering steel have demonstrated that the only Category A details are significantly affected (Albrecht and Naeemi 1984, Albrecht et al. 1989).

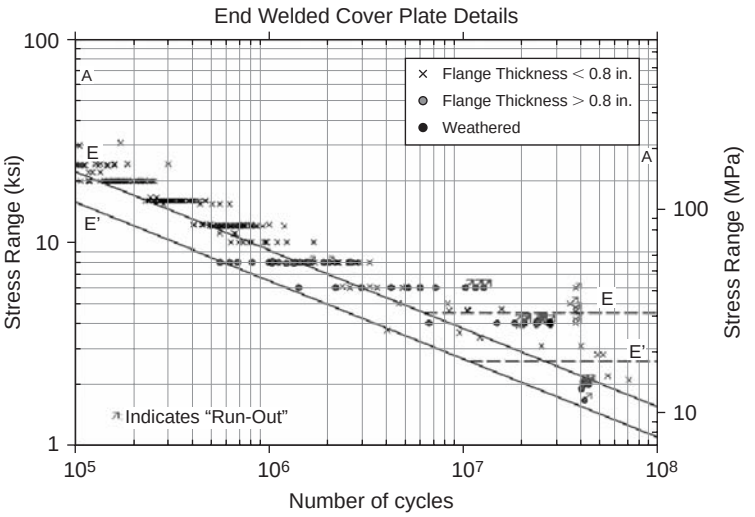


Figure 25 Test results on full-scale end welded cover plate beams.

7 High performance steels and enhancement

The development of high performance steels such as HPS Grade 485W and HPS Grade 690W has shown that without enhancement of welded details their fatigue resistance is no different than other high strength steels in use during the last four decades (Fisher et al. 1970). Post-weld enhancement of fatigue resistance of welded details such as cover plates, gussets and stiffeners that are known to experience crack growth from a weld toe is essential for an effective use of high strength modern steels. As noted in the previous section, peening and gas tungsten arc re melting has been used in the past to improve the fatigue resistance of welded details that have experienced fatigue damage.

Over the past decade, ultrasonic impact treatment (UIT) has proved to be a consistent and effective means of improving fatigue strength of welded connections. A number of investigators (Wright 1996, Statnikov 1997b, Haagensen et al. 1998) all reported that the fatigue strength of small size welded joints was improved by 50 to 200% at 2 million cycles when treated by UIT. Research at Lehigh University on large scale specimens having stiffener and cover plate welded details has also demonstrated that substantial increases in fatigue strength of these high strength steel welded details can be achieved by UIT in particular the elevation of their fatigue limit (Roy et al. 2003, Roy & Fisher 2005, 2006).

UIT involves post-weld deformation treatment of the weld toe by impacts from single or multiple indenting needles excited at ultrasonic frequency, generating force impulses at the weld toe (Statnikov 1997a). The treatment introduces beneficial compressive residual stresses at the weld toe and also reduces the stress concentration by enhancing the profile of the weld toe. The UIT equipment consists of a handheld tool consisting of an ultrasonic transducer, a wave guide, and a holder with impact needles; an electronic control box; and a water pump to cool the system. Compared with traditional impact

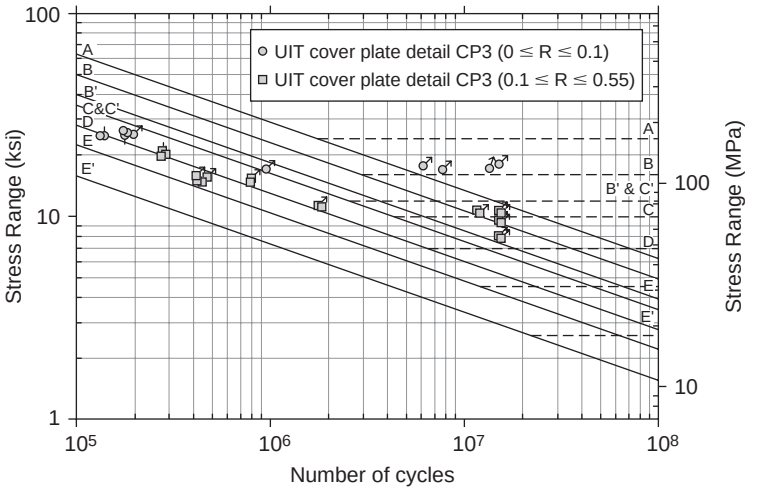


Figure 26 Fatigue performance of end welded cover plate details having weld size same as the plate thickness.

treatment methods such as air hammer peening, shot peening and needle peening, UIT appears to be more efficient and environmentally acceptable. It involves a complex effect of strain hardening, reduction in weld strain, relaxation in residual stress, and reduction in the stress concentration from profiling (Statnikov 1997a).

The large scale beam tests showed that although the treated details suppressed crack growth from the weld toe, when the usual end weld size was used the failure mode changed to fatigue crack growth from the weld root. This usually resulted in a longer life but still led to cracking and failure (Roy et al. 2003, Roy & Fisher 2005). For enhanced fatigue resistance it was desirable to prevent root cracking and this was achieved by increasing the size of the end weld at the cover plate to the plate thickness, which reduced the stress concentration at the weld root (Takamori & Fisher 2000). This concept was verified by the subsequent fatigue tests as can be seen from the experimental results plotted in Figure 26. These results showed that the enhancement in fatigue resistance was dependent on both the stress range S_r and the minimum stress S_{min}. A substantial improvement was realized at the lower level of minimum stress. This improvement was reduced when subjected to higher levels of minimum stress that was applied after the treatment. This characteristic is typical of other improvement techniques that introduce compression residual stress through plastic deformation (Fisher et al. 1979, Roy & Fisher 2006). Under low levels of minimum stress, the residual compression stress at the treated weld toe is more effective in suppressing crack growth as the discontinuities are not opened until a higher stress range is applied. At higher levels of minimum stress this condition is reached at a lower level of stress range causing a reduction in the level of enhancement and a lower level for the fatigue limit. Substantial enhancement results when the treatment is applied under a high level of minimum stress. The treatment is effective in reducing both the residual tensile stress from welding as well as the tension from the applied gravity load. This was verified experimentally for weld toes treated by air hammer peening (Fisher et al. 1979).

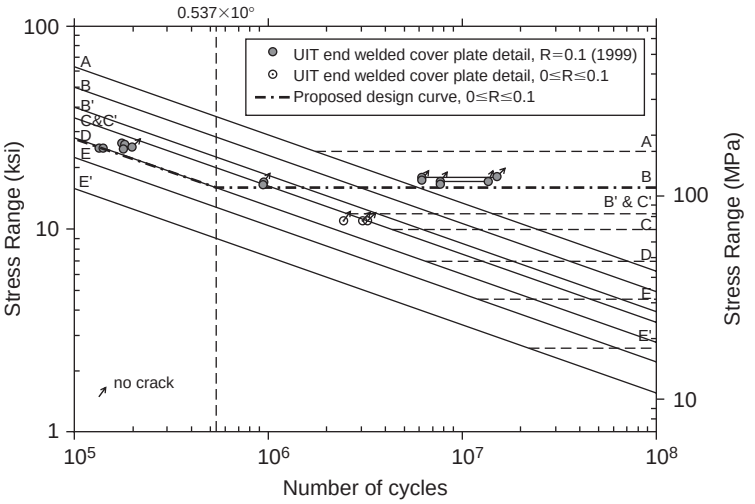


Figure 27 Design curve for end welded cover plate details treated by UIT; $0 \leq R \leq 0.1$.

Design curves for Category E′ end welded cover plates treated by UIT are provided in Figures 27 & 28. As is apparent in Figure 27, the treated details under low minimum stress (i.e. the R-ratio of S_{min}/S_{max} is less than 0.1 or the detail was treated under dead load) provide a design fatigue limit corresponding to Category B of the AASHTO specifications. None of the end-welded cover plate details developed fatigue cracks below 110 MPa, the CAFL for Category B (Roy & Fisher 2006). When the treatment is applied before the dead load stress is activated, and the R-ratio is bracketed by $0.1 < R < 0.5$, the design fatigue limit is decreased as shown in Figure 28 to 70 MPa which is the fatigue limit for Category C. Although there is enhancement in the finite life, it is not as significant as the increase in the fatigue limit for this class of detail.

Cover plate end welds on existing bridges are not likely to have end weld size same as the plate thickness. More likely the weld size will be about half the plate thickness. The test results indicated that when the R-ratio was less than 0.1, the enhanced fatigue resistance was applicable to the weld toe. There is a high probability that fatigue crack growth will initiate at the weld root as was demonstrated in the test girders removed from the Yellow Mill Pond bridge that were treated by air hammer peening and gas tungsten arc re-melting (Takamori & Fisher 2000). This would indicate that inspections should focus on the weld throat to ascertain if root cracking would subsequently develop. Fortunately, there is a significant increase in life for root cracking to occur and the cycles (time) necessary for the crack to propagate across the cover plate end to the longitudinal welds, which is the only way the crack can enter into the girder flange. Normal periods of inspection should identify such throat cracking if it ever occurs.

8 Conclusion

Fatigue of steel bridges under traffic loading is the most significant issue affecting the service performance of aging transportation infrastructure in the USA and in many

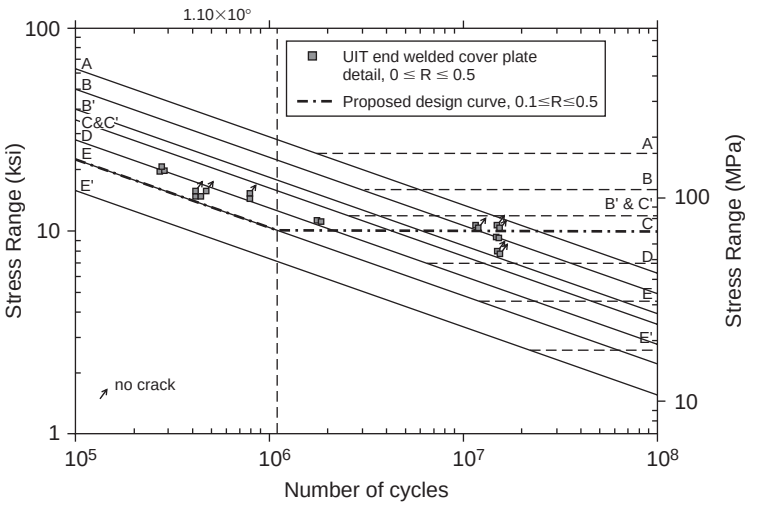


Figure 28 Design curve for end welded cover plate details treated by UIT; $0.1 < R \leq 0.5$.

other countries around the world. Research and case studies of in-service fatigue cracking of steel bridges over the past 40 years have helped in formulating design guidelines and improved detailing practices, implementation of which have limited fatigue cracking in new construction. However, the risk of fatigue fracture of many steel bridges that were built prior to the implementation of the current practices and the economic impact of replacing the deficient bridges remain high.

Historically most of the fatigue cracking of the welded steel bridges in the USA occurred at cover plate and similar attachment details, as well as at the web gaps from distortion. The attachment details are the most severe of the fatigue critical details, which are characterized by crack growth at the weld toe. While fatigue fracture limit state in new steel bridges can be suppressed by avoiding the fatigue critical Category D, E or E′ attachments, the performance of these details in existing bridges may be enhanced by weld toe treatments such as air hammer peening, GTA re-melting, or UIT. Post-weld toe treatments should also be considered in new structures for efficient use of modern HPS, where the attachment details including Category C′ connection plates and stiffeners cannot be avoided.

Distortion induced fatigue cracking in the web gaps may be solved by proper detailing that eliminates the secondary stresses driving these cracks. In most cases, the web-gap-cracking can be prevented by rigidly connecting the attachment plates to the tension flange. Where the distortion is displacement controlled, the stresses can be reduced by increasing the flexibility of the connection. If distortion is limited, holes may be drilled or cored at the crack tips to temporarily arrest propagation.

When the cumulative stress ranges in the variable stress spectrum exceeds the CAFL by 0.05% or more of the total stress cycles in the distribution, the fatigue resistance of the attachments is given by the extension of the linear sloped part of the S-N curve below the CAFL. An infinite life may be assumed when the cumulative exceedence of the stress cycles beyond the CAFL is limited to 0.01% of the total. Most structures

carry enough truck traffic to justify designing them for an infinite fatigue life, especially the deck elements.

The examples cited in this paper illustrate the role of corrosion phenomena in bridge service and the need to control the corrosion conditions on bridge structures. Proper inspection, cleaning of debris from corners and crevices, and avoidance of design features that promote corrosion grooving are all important steps for eliminating or reducing corrosion problems. Very extensive corrosion, such as that encountered on the Mianus River Bridge can create serious problems with the integrity of the structure. However, in most cases the general corrosion of bridge structures if monitored at reasonable intervals and corrected as necessary, should be considered a normal part of bridge service.

References

AASHTO 2004. *LRFD bridge design specifications*. Washington, D.C.: American Association of State Highway and Transportation Officials (AASHTO).

Albrecht, P.A. & Naeemi, A.H. 1984. Performance of weathering steel in bridges. *NCHRP Report 272*. Washington, D.C.: Transportation Research Board.

Albrecht, P.A., Coburn, F.M., Wattar, G.L., Tinklenberg, G.L. and Gallagher, W.P. 1989. Guidelines for the use of weathering steel in bridges. *NCHRP Report 314*. Washington, D.C.: Transportation Research Board.

AREMA 2005. *Manual of railway engineering*. Washington, D.C.: American Railway Engineering and Maintenance of Way Association (AREMA).

Connor, R.J. & Fisher, J.W. 2000. In-service response of an orthotropic steel deck compared with design assumptions. *Transportation Research Record* 1696(1): 100–108.

Connor, R.J. & Fisher, J.W. 2001. Results of field measurements on the Williamsburg Bridge orthotropic deck: Final report on Phase III. *ATLSS Report No. 01-01*. Bethlehem, PA: ATLSS Engineering Research Center, Lehigh University.

Connor, R.J., Richards, S.O. & Fisher, J.W. 2003. Long-term monitoring of prototype orthotropic deck panels on the Bronx-Whitestone Bridge for fatigue evaluation. In K.M. Mahmoud (ed.), *2003 New York City Bridge Conference; Proc.,* New York City, October 20–21. Lisse: Swets & Zeitlinger.

Demers, C.E. & Fisher, J.W. 1990. A survey of localized cracking in steel bridge structures, Vol. 1. *Report No. FHWA-RD-89-166*. Washington D.C.: Federal Highway Administration, U.S. Department of Transportation.

Dexter, R.J., Connor, R.J. & Kaczinski, M.R. 1997. Fatigue design of modular bridge expansion joints. *NCHRP Report 402*. Washington, D.C.: Transportation Research Board.

Fanjiang, G.-N., Ye, Q., Fernandez, O.N. & Taylor, L.R. 2004. Fatigue analysis and design of steel orthotropic deck for Bronx-Whitestone Bridge, New York City. *Transportation Research Record* 1892: 69–77.

Fisher, J.W., Frank, K.H., Hirt, M.A. & McNamee, B.M. 1970. Effect of weldments on the fatigue strength of steel beams. *NCHRP Report 102*. Washington, D.C.: Highway Research Board.

Fisher, J.W., Hausammann, H., Sullivan, M.D. & Pense, A.W. 1979. Detection and repair of fatigue damage in welded highway bridges. *NCHRP Report 206*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Barthelemy, B.M., Mertz, D.R. & Edinger, J.A. 1980. Fatigue behavior of full-scale welded bridge attachments. *NCHRP Report 227*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Mertz, D.R. & Zhong, A. 1983. Steel bridge members under variable amplitude long life fatigue loading. *NCHRP Report 267*. Washington, D.C.: Transportation Research Board.

Fisher, J.W. 1984. *Fatigue and fracture in steel bridges: case studies*: John Wiley.

Fisher, J.W. & Pense, A.W. 1984. Mianus River Bridge material: Fracture surfaces and special tests. *Final Report*. Zetlin-Argo Structural Investigations, Inc. Connecticut Department of Transportation.

Fisher, J.W., Yen, B.T., Wang, D.W. & Mann, J.E. 1987. Fatigue and fracture evaluation for rating riveted bridges. *NCHRP Report 302*. Washington DC.: Transportation Research Board.

Fisher, J.W., Jin, J., Wagner, D.C. & Yen, B.T. 1990. Distortion-induced fatigue cracking in steel bridges. *NCHRP Report 336*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Yen, B.T. & Wang, D. 1990. Fatigue strength of riveted bridge members. *Journal of Structural Engineering* 116(11).

Fisher, J.W., Nussbaumer, A., Keating, P.B. & Yen, B.T. 1993. Resistance of welded details under variable amplitude long-life fatigue loading. *NCHRP Report 354*. Washington, D.C.: Transportation Research Board.

Haagensen, P.J., Statnikov, E.S. & Lopez-Martinez, L. 1998. Introductory fatigue tests on welded joint in high strength steel and aluminum improved by various methods including ultrasonic impact treatment (UIT). *IIW Doc. No. XIII-1748-98*. Paris: International Institute of Welding.

Keating, P. & Fisher, J.W. 1986. Evaluation of fatigue tests and design criteria on welded details. *NCHRP Report 286*. Washington, D.C.: Transportation Research Board.

Kunihiro, T., Inoue, K. & Fukuda, T. 1972. Atmospheric exposure of weathering steel. *Report Br. 71-08*. Tokyo Research Laboratory, Ministry of Construction.

Moses, F., Schilling, C.G. & Raju, K.S. 1987. Fatigue evaluation procedures for steel bridges. *NCHRP Report 299*. Washington, D.C.: Transportation Research Board.

National Transportation Safety Board (NTSB). 1983. Highway Accident Report: Collapse of a suspended span of Interstate Route 95 Highway Bridge over the Mianus River, Greenwich, Connecticut. *Report No NTSB/HAR-84/03*.

Roy, S., Fisher, J.W. & Yen, B.T. 2003. Fatigue resistance of welded details enhanced by ultrasonic impact treatment (UIT). *International Journal of Fatigue* 25(9–11): 1239–1247.

Roy, S. & Fisher, J.W. 2005. Enhancing fatigue strength by ultrasonic impact treatment. *International Journal of Steel Structures* 5(3): 241–252.

Roy, S. & Fisher, J.W. 2006. Modified AASHTO design S-N curves for post-weld treated welded details. *Journal of Bridge Structures – Assessment, Design and Construction* 2(4): 207–222.

Schilling, C.G., Klippstein, K.H., Barsom, J.M. & Blake, G.T. 1978. Fatigue of welded steel bridge members under variable amplitude loading. *NCHRP Report 188*. Washington, D.C.: Transportation Research Board.

Shastry, C.R., Friel, J.T. & Townsend, H.R. 1987. Sixteen-year atmospheric corrosion performance of weathering steels in marine, rural and industrial environments. In S.W. Dean and T.S. Lee (eds.), *Symposium on Degradation of Metals in the Atmosphere, Proc., ASTM STP 965*. Philadelphia, PA: ASTM International.

Sivakumar, B., Khoury, E., Fisher, J.W. & Wright, W.J. 2003. Hoan bridge brittle fracture and retrofit design. In K.M. Mahmoud (ed.), *2003 New York City Bridge Conference; Proc.,* New York City, October 20–21. Lisse: Swets & Zeitlinger.

Statnikov, E.S. 1997a. Applications of operational ultrasonic impact treatment (UIT) technologies in production of welded joint. *IIW Doc. No. XII-1667-97*. Paris: International Institute of Welding.

Statnikov, E.S. 1997b. Comparison of post-weld deformation methods for increase in fatigue strength of welded joints. *IIW Doc. No. XII-1668-97*. Paris: International Institute of Welding.

Takamori, H. & Fisher, J.W. 2000. Tests of large girders treated to enhance fatigue strength. *Transportation Research Record* 1696: 93–99.

Tilly, G.P. & Nunn, D.E. 1980. Variable amplitude fatigue in relation to highway bridges. *Proc. the Institution of Mechanical Engineers (London)* 194: 259–267.

Wright, W.J. 1996. Post-weld treatment of a welded bridge girder by ultrasonic hammer peening. *FHWA Internal Research Report*. McLean, VA: Federal Highway Administration, Turner Fairbank Highway Research Center.

Wright, W.J., Fisher, J.W. & Kaufmann, E.J. 2003. Failure analysis of the Hoan Bridge fractures. In K.M. Mahmoud (ed.), *2003 New York City Bridge Conference; Proc.*, New York City, October 20–21. Lisse: Swets & Zeitlinger.

Yamada, K. & Kikuchi, Y. 1984. Fatigue tests of weathered welded joints. *ASCE Journal of the Structural Division* 110(9).

Chapter 3

Prediction of reinforcement corrosion in concrete structures

A.G. Razaqpur

Centre for Effective Design of Structures, McMaster University, Department of Civil Engineering, Hamilton, Ontario, Canada

O.B. Isgor

Carleton University, Department of Civil and Environmental Engineering, Ottawa, Ontario, Canada

ABSTRACT: This paper deals with the problem of predicting the short- and long-term corrosion rate in reinforced concrete structures. The paper briefly describes the fundamental concepts of electrochemistry as it pertains to corrosion of reinforcement in concrete and the effect of the various factors which contribute to the incidence and rate of corrosion. This is followed by a discussion of the available field measurement techniques for predicting corrosion incidence and rate, and the difficulties involved in interpreting the field data. Next some prominent empirical, theoretical and numerical models of corrosion prediction are reviewed and their advantages and disadvantages are briefly described. It is contented that a combination of field measurements and numerical models may be the best way to obtain a reasonable estimate of corrosion activity in a structure over the long-term. The latter is demonstrated through a case study involving a building with chloride contamination.

1 Introduction

Corrosion of steel reinforcement in exposed concrete structures is one of the most prevalent causes of their deterioration and it adversely affects their serviceability and safety. Generally, corrosion occurs when the alkalinity of concrete is lowered by the intrusion of carbon dioxide into concrete, termed carbonation, or by the diffusion and buildup of chlorides adjacent to the film. Carbonation is due to the presence of high concentrations of carbon dioxide in the atmosphere produced by burning of fossil fuels, which may be more severe in urban environments with high automobile emissions over a long period. On the other hand, chlorides enter concrete structures when they are exposed to deicing salts or to marine environments. Chlorides may also be part of the concrete mix as contaminant in fine and coarse aggregates or as additive to accelerate setting of the concrete.

Although the causes for the incidence of corrosion are well known, it is not always easy to completely avoid the circumstances that induce reinforcement corrosion. Therefore, engineers are interested in knowing the time to the initiation and subsequent rate of propagation of corrosion because corrosion propagation is accompanied by damage to the concrete and steel reinforcement, which impair the serviceability and safety of structures. Determination of the rate of corrosion of steel in concrete structures is of critical importance for assessing their durability and long-term performance, and for

prioritizing maintenance and repair budgets and operations. The corrosion rate can be estimated using either nondestructive field measurement techniques or can be computed using various empirical and numerical techniques. Although it may seem logical to assume that actual field measurements would provide a more accurate assessment of the extent of corrosion and its effect on the safety of structures, in fact this may not be always the case, unless an exorbitant amount of time and effort are expended, which may render the exercise expensive, time consuming and impractical. As will be explained later, the corrosion current measured in the field cannot be used to determine with certainty the number of bars or the fraction of the total surface area of any bar contributing to the measured corrosion current. It is particularly difficult to identify pitting corrosion by measuring corrosion current.

Empirical and numerical methods are similarly afflicted by many uncertainties. Empirical methods are invariably derived from either control experiments, which focus on a limited number and range of parameters, or from field measurements which often pertain to specific environmental and structural conditions, the extrapolation of which to other conditions is fraught with much uncertainty. Advanced numerical techniques which include precise heat and moisture transfer and modeling of the electrochemistry of corrosion require specification of both initial and boundary conditions, knowledge of the corrosion potential of reinforcement and the electrochemical and physical properties of concrete. Often quantitative information about these properties is lacking and it is particularly difficult to forecast their variation over long time periods. On the other hand, advanced numerical simulations can be performed assuming certain plausible scenarios and the expected performance of the structure can be circumscribed within the confines of the assumptions made. This procedure would allow one to evaluate the consequences of the best and worst anticipated conditions. Ideally, numerical simulations should be coupled with field measurements to arrive at a reasonable assessment of the future performance of the structure.

In this paper existing corrosion measurement techniques, and empirical and numerical techniques for predicting corrosion rate and amount are reviewed and their advantages and disadvantages are highlighted. It is asserted that a combination of field measurements and advanced numerical modeling may yield the most reliable prediction of corrosion rate. The latter is demonstrated by an example from actual practice. In order to facilitate the discussion, we begin with the description of the electrochemistry of corrosion and the significance of the various environmental, physical and chemical factors that affect the incidence and rate of corrosion.

2 Fundamentals of corrosion of steel in concrete

2.1 *Corrosion initiation*

It is common practice to divide the corrosion process in concrete structures into two distinct phases. Phase one, termed the initiation period, refers to the time required to reach the necessary conditions in the concrete for the reinforcement to begin active corrosion. Phase two, called propagation period, commences at the end of the initiation period and continues as long as the reinforcement is experiencing corrosion. Within the high alkaline environment of concrete (pH >12.5), reinforcing steel is covered with a thermodynamically stable film of iron oxides which is insoluble in the concrete pore

solution and protects the steel against further corrosion. The partial or complete loss of this film, termed depassivation, deprives steel of its protective coating and renders it vulnerable to corrosion. The high alkalinity of the concrete pore solution is due to the presence of sodium, potassium and calcium hydroxide, which are products of the hydration of cement. The lowering of the concrete pH at steel-concrete interface by carbonation and/or the accumulation of chlorides at the same location creates the thermodynamic conditions for the dissolution of the protective oxide film on the steel surface.

When carbon dioxide from the atmosphere diffuses into concrete, it forms an acidic solution which reacts with the alkaline solution in the concrete pores and lowers its pH. The chemistry of this process is represented by the following reactions:

(a) Formation of acidic solution:

$$CO_2 + H_2O \rightarrow 2H^+ + CO_3^{2-} \tag{1}$$

(b) Neutralization of concrete pore solution:

$$2H^+ + CO_3^{2-} + Ca^{2+} + 2OH^- \rightarrow CaCO_3 + 2H_2O \tag{2}$$

Reaction [2] may be followed by the following reaction:

$$CaCO_3 + H_2O + CO_2 \rightarrow Ca(HCO_3)_2 \tag{3}$$

Reaction [2] indicates that carbonation removes the hydroxyl ions from the pore solution, thus lowering its pH. Once the pH falls below approximately 10, the passive film becomes unstable and breaks down. Since carbonation reactions involve water, the concrete humidity must be sufficiently high for this reaction to proceed. On the other hand, for carbon dioxide as a gas to enter concrete, some pores must be partially empty. Hence, neither fully saturated concrete nor concrete with low moisture content is susceptible to carbonation. A relative humidity in the range of 50% to 70% is most conducive to carbonation (Neville, 1995). Once the carbonation front reaches the steel-concrete interface, corrosion will normally commence. Since carbon dioxide moves by diffusion through the concrete pore solution, the greater the reinforcement cover thickness and the denser the concrete microstructure, the longer it will take for the carbonation front to penetrate to the reinforcement level.

Chloride attack generally occurs when the concentration of chlorides in the concrete adjacent to the reinforcement, often expressed in terms of percent by mass of cement or binder, exceeds a certain threshold or limit. The threshold depends on the pH of the concrete, the corrosion potential of the steel and the presence of voids in the concrete in the vicinity of the reinforcement (Bertolini et al., 2004). For instance, Bertolini et al. (2004) report that in structures exposed to the atmosphere the threshold is 0.4%–1% of the cement content while for structures immersed in water it is much higher. In the former case, the corrosion risk associated with concentrations below 0.2% is deemed low while that associated with concentrations greater than 1.0% is deemed high. Note that part of the chlorides in concrete are chemically bound by the C_3A phase in the hydrated cement or adsorbed on the surface of the pore walls, with the remaining amount being free in the pore solution. It is the free chlorides that can attack

the passive film albeit the proportion of the free and the bound chlorides for a given total amount of chlorides in concrete can easily change due to changes in the chemical and physical conditions of the pore solution. Nevertheless, the ability of concrete to bind chlorides will increase its corrosion protection ability. Poor quality concrete with high water/cement (w/c) ratio, or concrete with many wide cracks, would enable rapid ingress of chlorides and carbon dioxide inside the concrete, which would lead to premature corrosion and deterioration of the structure. On the contrary, structure made of a high quality concrete with a uniform and dense microstructure and possessing a thick cover over the reinforcement would impede the ingress of carbon dioxide and chlorides and would consequently prolong the time to the initiation of the corrosion process. Other factors which will affect the depassivation of steel in concrete include the surface characteristics and chemical composition of steel, and sustained mechanical stresses (Uhlig and Revie, 1985; Neville, 1995; Broomfield, 1997). It has been suggested that $CaCl_2$ also causes a reduction in pH of concrete while NaCl does not affect pH significantly. Consequently, deicing salts containing mainly calcium chloride are more aggressive than those containing primarily sodium chloride.

The corrosion of steel reinforcement can occur in the form of pitting corrosion or uniform corrosion. Pitting may occur where the chloride ion concentration locally around the steel reinforcement exceeds the threshold value. In the case of the carbonation of concrete cover, the corrosion generally takes place uniformly along the length of the bar or over the surface of an entire reinforcement mat. The external manifestation of general or uniform corrosion is extensive cracking and delamination of the concrete cover. On the other hand, extensive pitting corrosion can cause locally major loss of rebar cross-section without noticeable surface damage to concrete. Depending on the available to the required capacity of the reinforcement at the affected section, this form of corrosion could rapidly endanger the safety of a structure.

2.2 *Corrosion propagation*

Once the steel is depassivated, corrosion will begin and propagate. Corrosion is an electrochemical reaction, which involves anodic and cathodic half-cell reactions. At the anode, the iron in the steel dissolves into the concrete pore solution according to the following half-cell reaction

$$Fe \rightarrow Fe^{2+} + 2e^{-} \tag{4}$$

The electrons released by the anodic reaction are consumed at the cathodic sites on the steel surface to preserve electrical neutrality. The cathodic reaction is the oxygen reduction according to

$$O_2 + 2H_2O + 4e^{-} \rightarrow 4(OH^{-}) \tag{5}$$

If the anodic reaction given in reaction [4] were the only reaction that took place during the corrosion process, there would be no cracking and spalling of the concrete surrounding the steel. However, Fe^{2+} ions react within the pore solution to form the

common rust. The following reactions represent the formation of the so-called "red rust" after iron dissolution occurs at the anodic sites (Broomfield, 1997):

$$Fe^{2+} + 2OH^- \rightarrow Fe(OH)_2 \tag{6}$$

$$4Fe(OH)_2 + 2H_2O + O_2 \rightarrow 4Fe(OH)_3 \tag{7}$$

$$2Fe(OH)_3 \rightarrow 2H_2O + Fe_2O_3 \cdot H_2O \tag{8}$$

$Fe_2O_3 \cdot H_2O$, or red rust, is the main corrosion product, but the other iron oxides may also exist. Therefore, the volume of the corrosion products is three to four times larger than the volume of the original steel (iron) involved in the corrosion reaction. The various iron oxides have volume 2 to 6 times that of the iron in them. Due to its large volume, the formation of red rust causes internal stresses and generally leads to the cracking and spalling of the concrete surrounding the reinforcement. In the case of pitting corrosion, the iron oxides may simply diffuse away from the corrosion site and appear as stain on the concrete surface without causing cracking. Although additional reactions other than the ones given in reactions [4] to [8] may take place during the corrosion process, (e.g., hydrogen evolution due to water decomposition at extremely active locations on the steel surface, or the formation of black rust), in the present study, the corrosion of steel in concrete will be defined by the reactions [4] to [8]. Notice that both the anodic and cathodic reactions involve water and oxygen, thus for the reactions to proceed, water and oxygen must be available in the concrete pores near the steel surface. Oxygen in these reactions needs to be in the dissolved state in the pore solution.

Oxygen enters concrete as a gas, and similarly to carbon dioxide, it requires some empty pore space to diffuse. Hence oxygen cannot easily enter fully saturated concrete and therefore negligible corrosion occurs in concrete structures fully immersed in water. On the other hand, since water is an essential part of the corrosion reaction, it must be available within the pore solution. Consequently, the highest corrosion rate may occur in concrete with a relative humidity in the range of 95 to 98%. The rate of corrosion in nearly saturated concrete structures is to a large extent controlled by the rate of diffusion and availability of oxygen at the cathodic site.

2.3 *Corrosion potential*

The half-cell reactions [4] and [5] each have a certain electric potential, known as the standard potential. This potential can be found relative to a reference or standard electrode using an electrochemical cell. Note that the measured potential would be relative to the standard electrode and in practice various types of standard electrodes are used, including hydrogen (SHE), copper/copper sulfate (CSE), silver/silver chloride (SSE), and standard calomel (SCE). Thus, the standard electric potential of an electrode can have different values, depending on the reference electrode. Relative to the hydrogen electrode, the standard electrode potential of half-cell reactions [4] and [5] are $E_a^0 = +0.440$ V (Volt) and $E_c^0 = +0.401$ V, respectively. The subscripts a and c refer to the anode and cathode while superscript 0 denotes standard conditions. The difference between the cathodic and anodic potentials constitutes the driving force for

corrosion. Since this potential is negative, corrosion could occur without an external driving force. In other words, the more negative the cell potential, the greater the probability of corrosion. This is the theory behind the ASTM C876 (1991) half-cell potential measurement method for determining the corrosion potential in a reinforced concrete structure. According to this method, using a copper-copper sulfate standard electrode, if the measured half-cell potential is more positive than −200 mV, then there is a greater than 90% probability that no corrosion is occurring while if it is more negative than −350 mV, then there is a greater than 90% probability of corrosion occurring. For potential values between those limits, corrosion may or may not be occurring.

In practice, the corrosion cell temperature and electrolyte concentration will vary from the standard conditions, thus the electrode potential will change, but it can be determined using Nernst equation (Uhlig and Revie, 1985). The anodic and cathodic electric potentials at a temperature of 25°C and different concentrations of iron ions and oxygen can be written as (Ahmad, 2003)

$$E_a(V) = -0.44 + 0.0296 \log [Fe^{2+}] \tag{9}$$

$$E_c(V) = 1.229 + 0.0148 \log [O_2] - 0.0591 \text{ pH} \tag{10}$$

Consequently, the electromotive force deriving this cell can be written as:

$$\Delta E(V) = E_c - E_a = 1.229 + 0.0148 \log [O_2] - 0.0591 \text{ pH} - 0.0296 \log [F\ e^{2+}] \tag{11}$$

The electromotive force generates an electric current. Once a current is generated, it alters the cathodic and anodic potentials through a phenomenon known as polarization. The polarization lowers the potential of the passive steel (cathodic sites) and increases that of the active steel (anodic sites), Fig. 1.

Thus part of the driving force in Eq. 11 is consumed to polarize the anode and the cathode. Since the corrosion current involves the movement of ions in the concrete pore solution, and concrete is not perfectly conductive, the electromotive force must also overcome the electrical resistance of concrete, a phenomenon known as ohmic drop. The electrical resistance of concrete depends on its moisture content; the higher the moisture content, the lower the resistance of concrete. Once concrete dries out, its resistance increases several orders of magnitude and a large electromotive force is needed to drive a current through it. In drier concrete, only the concrete in the immediate neighborhood of the actively corroding bars may serve as cathodic site. Thus the rate of corrosion will be controlled by the availability of oxygen over this small area. On the other hand, in concrete with high humidity, a large portion of the concrete sufficiently remote from the anodic site can serve as cathode and therefore extensive corrosion can occur even if the oxygen in some areas is depleted. Generally, in reinforced concrete the rate of corrosion is determined by the rate of cathodic reaction, which depends on the availability of oxygen and water. Based on the above, and assuming a simplified equivalent electric circuit (similar to the one used by Raupach and Gulikers as illustrated in Fig. 2), the corrosion current density, i_{corr} (A/m^2) may be

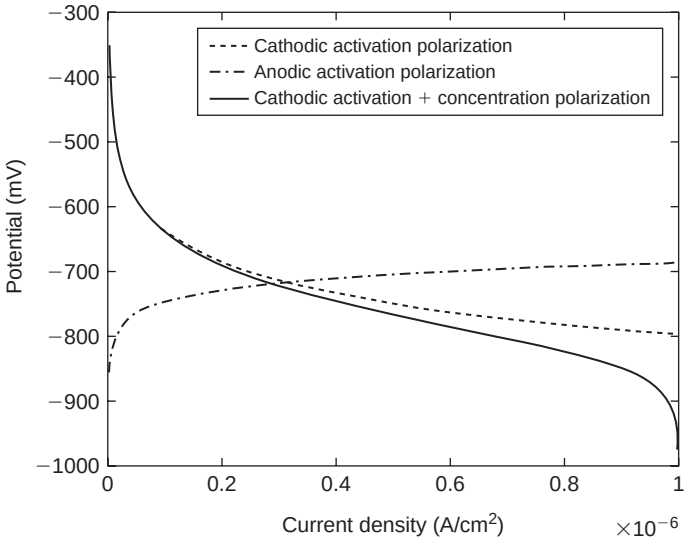


Figure 1 The polarization of anodes and cathodes on steel surface (Ge, 2006).

expressed as (Hunkeler, 2005)

$$i_{corr} = \frac{I_{corr}}{A_a} = \frac{\Delta E}{A_a(R_a + R_c + R_{oh})} = \frac{\Delta E}{r_a + \frac{r_c A_a}{A_c} + \frac{\rho A_a}{L}} \tag{12}$$

where R = resistance (Ω), A = area (m^2), r = specific resistivity = $R \cdot A$ ($\Omega \cdot m^2$), ρ = specific resistivity of concrete ($\Omega \cdot m$), L = characteristic length (m) and subscript a and c denote anodic and cathodic. Note that the corrosion current depends on the ratio of the anodic to the cathodic areas. As the cathodic area decreases relative to the anodic area, the denominator of Eq. 12 increases and the corrosion current density decreases.

2.4 *Corrosion rate*

The rate of rust production on the steel surface, J_{rust} [kg/m^2/s], is related to the corrosion current density by Faraday's law. The rate of ferrous oxide, $Fe(OH)_2$, production, J_{Fa}, at the anodic regions can be written as (Bazant, 1979):

$$J_{Fa} = \frac{i_a}{zF} = 4.656 \times 10^{-7} i_a \tag{13}$$

where i_a is the anodic current density, F is the Faraday's constant (9.65×10^4 C/mol), and z is the number of electrons exchanged in the reaction ($z = 2$ for steel corrosion). $Fe(OH)_2$ can be further oxidized, and this will result in the production of $Fe(OH)_3$. Since one mole of $Fe(OH)_2$, which is 89.845 g, produces one mole of $Fe(OH)_3$

(106.845 g), the rate of rust production, J_{rust}, at the anodic regions on the steel surface can be calculated by:

$$J_{rust} = \frac{106.845}{89.845} J_{Fa} = 5.536 \times 10^{-7} i_a \tag{14}$$

Alternatively, one could relate the corrosion current density, expressed in A/m^2, to the rate of steel thickness loss, often called the corrosion penetration rate, p (mm/year), as follows (Gulikers, 2005):

$$p = 31.56 \times 10^6 \frac{W i_{corr}}{zFG} \tag{15}$$

in which W = atomic weight of iron = 56 g/mole, and G = density of iron = 7874 kg/m^3.

Therefore, once the current density at any point along the steel bars is calculated, the determination of the rate and amount of corrosion becomes a straightforward task. For example, a current density of 1.0 A/m^2 is equivalent to a corrosion penetration rate of 1.16 mm/year, which is equivalent to a mass loss of approximately 9.125 kg/m^2 per year. Corrosion penetration up to 100 μm in ordinary reinforced concrete structures is deemed acceptable. If the design life of the structure were 75 years, and the corrosion initiation period was 25 years, this would translate into an average penetration rate of 2.0 μm/year.

It should be noted that conversion of the corrosion current to the corrosion current density requires knowledge of the actual surface areas of the anodic and cathodic regions on the steel surface. However, to date no simple method exists that can quantify these areas. Consequently, the main difficulty in this process is the calculation of current density in the steel. It is this difficulty that partly makes the interpretation of both field measurements and computational methods complicated and can lead to erroneous conclusions. Notwithstanding this difficulty, numerous models exist for predicting corrosion and various field measurement techniques have been developed to estimate it. With this problem in mind, in the following section, a review of some corrosion measurement techniques and corrosion rate prediction models is presented.

3 Corrosion measurement methods

One of the most widely used methods for measuring corrosion rate in concrete structures is to measure the polarization resistance of the corroding system and insert it in the Stern-Geary equation to determine the corrosion current density, i_{corr} (A/m^2) (Elsener, 2005)

$$i_{corr} = \frac{B}{R_p} \tag{16}$$

where B = constant (V) and R_p = polarization resistance (Ω^{-1} m^{-2}). The constant B is normally 26 mV for corroding and 52 mV for non-corroding steel. There are a number of corrosion measurement devices that use this approach (Elsener, 2005), but in each case, the calculations that are based on measured polarization resistance data

provides instantaneous corrosion rates that are influenced greatly by prevailing temperature and humidity; therefore comparisons with average corrosion rates, which can be determined through weight loss measurements, can only be carried out after integrating polarization resistance data over time. In order to accurately predict the corrosion rates in concrete, one needs to monitor the polarization resistance frequently for a period of time so that temporal variations can be captured. Therefore, although the polarization resistance methods are quite practical for instantaneous monitoring of structures, they become significantly expensive and impractical if the monitoring is to be done continuously over long periods.

In addition to the practical disadvantages of using instantaneous corrosion rate measurement techniques that are based on polarization resistance data, there are also inherent difficulties associated with the procedure. The polarization resistance techniques use auxiliary tools (e.g. guard rings) that limit the polarized surface area of the reinforcement to obtain the polarization resistance data from this polarized surface. It has been demonstrated (Pour-Ghaz, 2007) that limiting the polarized surface area (using guard rings) can be a challenge; typically the actual polarized surface area is larger than the assumed area. Even when guard rings work well, and the desired surface of the reinforcement is polarized, there is an inherent problem with the approach due to the fact that the polarization resistance technique normalizes the measured corrosion rate over the entire polarized area of the reinforcing steel. In the case of pitting corrosion, since the corroding area (i.e. anode) can be significantly smaller than the polarized length of the steel, these instruments may underestimate the corrosion rate (Pour-Ghaz, 2007; Elsener 2002). It has been reported in the literature that corrosion rates measured by instruments that are based on polarization resistance technique can underestimate the actual corrosion rates by as much as 10 times (Luping, 2002).

Another widely used, practical, and standardized non-destructive method for monitoring steel corrosion in concrete structures is half-cell potential mapping (ASTM C876-91, RILEM TC 154-EMC). In this method, the potential difference between an external electrode located at the surface of concrete and the embedded reinforcement is measured with a high impedance voltmeter. A solid electrical connection between the reinforcement and the voltmeter, and a wet connection between the external electrode and the reinforcement through the concrete cover are essential for obtaining reliable readings. Despite its widespread application, half-cell potential mapping is usually associated with a number of practical difficulties which can be summarized as follows (Berkeley and Pathmanaban, 1990): (1) Establishing a solid connection to the reinforcement, particularly in the case of densely reinforced members, such as bridge decks, can be a difficult task; (2) Establishing a proper wet connection between the external electrode and the reinforcement through the concrete cover can be problematic since a moisture (hence, resistivity) gradient always exists in the bulk of concrete, affecting the half-cell potential readings; (3) The time required to establish an equilibrium condition between the external electrode and concrete is a function of a number of parameters, including the thickness and the composition of the concrete cover (RILEM TC 154-EMC). Furthermore, half-cell potential mapping only provides information about the probability of corrosion occurrence, and it does not give any insight about the rate (or kinetics) of corrosion.

It should also be noted that half-cell potential measurement is also a function of the type of corrosion process. In the case of uniform corrosion, the potential readings at

the surface of concrete are generally close to the potential at the interface of steel and concrete (Sagues and Kranc, 1992); in the case of non-uniform corrosion, however, the measured potentials at the surface of concrete can be substantially different from those at the steel/concrete interface. This potential difference is a function of cover thickness and concrete resistivity, and increases with both. Several methods for compensating this potential difference are available in the literature (Uhlig and Revie 1985); however, the majority of these methods are not applicable to half-cell potential measurements of non-uniform corrosion of steel reinforcement.

Although some of the difficulties associated with polarization resistance and half-cell measurements can be minimized with the aid of experienced staff and accurate instrumentation, the interpretation of the results remains to be a major challenge for engineers. Concrete resistivity, oxygen availability, anode–to-cathode area ratio (A_a/A_c), and cover thickness are some of the factors that can influence the test (Pour-Ghaz, 2007; Elsener, 2002, Bertolini, et al., 2004); therefore, the results of the tests should be interpreted considering these factors.

4 Corrosion modelling

Considering the above difficulties, accurate modeling of steel corrosion in concrete structures is an important tool that can help in the better interpretation of the data from corrosion measurement techniques. Although it is difficult, if not impossible, to completely replace the nondestructive testing methods with theoretical models, accurate simulations combined with relatively infrequent, and hence inexpensive, polarization resistance measurements can be the answer to health monitoring of reinforced concrete infrastructure. These models can be classified under two categories: empirical and mathematical.

Empirical models are based on observed correlation between corrosion rate of steel in concrete and different parameters affecting it. Theoretical models are based on either the mathematical representation of the controlling mechanisms of corrosion or the detailed numerical modeling of polarization of steel. In this paper, a number of empirical and mathematical models are presented; for a detailed review of a more extensive list of models reference can be made to Pour-Ghaz (2007).

4.1 *Empirical models*

Empirical models are based on the observed relationship between the corrosion rate of steel and different parameters affecting it such as concrete properties and environmental conditions. These models are easy to use, but because they are derived from limited experimental data and because they do not consider the fundamental mechanisms of corrosion from a theoretical point of view, their general applicability and reliability may be questionable. In the following sections, a number of prominent empirical models for the determination of steel corrosion rate in concrete are presented without providing details.

4.1.1 *Morinaga (1990)*

An empirical model for predicting chloride-induced steel corrosion rate was developed by Morinaga (1990). Chloride contaminated specimens were exposed to outdoor

environment for a 10-year period, and the corrosion rate of steel in each specimen was determined by mass loss method. The following relationship was developed by regression analysis using the results of the experiments:

$$i_{corr} = [-0.51 - 7.60C_{Cl} + 44.97(w/c)^2 + 67.95C_{Cl}(w/c)^2] \times (d_{st}/d^2) \quad (17)$$

where i_{corr} in this equation was represented in terms of 10^{-4}g/cm^2/year; C_{Cl} (% by weight of mixing water) is the chloride content in the form of NaCl; d_{st} (mm) is the diameter of reinforcing steel and d (mm) is the concrete cover thickness.

Morinaga's model can be considered to be quite comprehensive since they are built on long-term studies that cover a large number of test variables. Nevertheless, in the case of chloride-induced corrosion, the design and environmental parameters were separated from each other, and no guideline was provided for merging them. In practice, these parameters cannot be considered separately. It should also be noted that for investigating the environmental effects on the chloride-induced corrosion, Morinaga used grout coated rebars instead of concrete or mortar specimens. Extrapolating the results from grout coated specimens to concrete and mortar is prone to error since many properties (e.g. porosity, water content, diffusivity) of grout coatings are substantially different from those of the concrete and mortar.

4.1.2 *Liu and Weyers (1998)*

Liu and Weyers (1998) developed an empirical model by testing large reinforced concrete slabs specimens simulating concrete bridge decks. A total of 44 specimens were produced with different values of chloride content and cover thickness. The w/c ratio of the specimens varied between 0.41 and 0.45. Corrosion rate and electrical resistance of the specimens were measured with 3-Electrode-Polarizarion (3LP) device. Temperature was monitored with a T-type thermocouple embedded at rebar depth. At the final stage of the experiment, metal mass loss measurements were performed in accordance with ASTM G1-90 method C3.5. During a 5-year outdoor exposure of the specimens, a total of 2927 measurements were recorded. The following relationship was developed by a regression analysis on the data recorded by 3LP.

$$\ln 1.08i_{corr} = 7.89 + 0.7771 \ln 1.69C_t - \frac{3006}{T} - 0.000116R_{con} + 2.24t^{-0.215} \quad (18)$$

where C_t (kg/m^3 of cement) is acid soluble chloride content as determined by ASTM C1152, T (°K) is temperature, R_{con} (Ω) is the resistance of concrete, and time, t, is given in years.

In addition to the fact that the applicability of the Liu and Weyers model is limited by the range of experimental parameters that it was obtained from, it also does not include the effect of oxygen availability on the steel surface; in other words, it does not include concentration polarization. In addition, since resistance, instead of resistivity, is used in the model, the predictions of the model are geometry dependent.

4.1.3 *DuraCrete (2000)*

Since estimating the corrosion rate of steel reinforcement by measuring the electrical resistance of concrete is convenient, there has been growing interest in developing models that are based on the relationship between corrosion rate and concrete resistance. As part of the European DuraCrete project (2000) the following empirical equation for predicting corrosion rate was proposed:

$$i_{corr} = \frac{10^4}{\rho(t)} F_{Cl} F_{Galv} F_{Oxide} F_{O_2} \tag{19}$$

where ρ ($\Omega \cdot$ m) is the concrete resistivity, F_{Cl} is the chloride content factor, F_{Galv} is the galvanic effect factor, F_{Oxide} is a factor related to the presence of oxide layer at the surface of steel, and F_{O_2} is the oxygen availability factor. The concrete resistivity at time t, $\rho(t)$ ($\Omega \cdot$ m), in Eq. 19 is given as:

$$\rho(t) = \rho(t_o) f_e f_t \left(\frac{t}{t_o}\right)^{n_{age}} \tag{20}$$

where $\rho(t_o)$ ($\Omega \cdot$ m) is concrete resistivity measured at time t_o, f_e is the exposure condition factor, f_t is the test method factor, and n_{age} is age constant.

It should be noted that in the DuraCrete model, there is a lack of data to quantify the introduced factors and their interactions (Raupach 2006). Furthermore, this model is developed only for chloride-induced corrosion; therefore the corrosion rate of carbonation-induced corrosion may not be estimated accurately since chloride contamination reduces the concrete resistivity whereas carbonation increases it. In addition, DuraCrete model only considers the resistance control mechanism and ignores all the other control mechanisms such as diffusion and activation. This causes the model to be limited to a certain relative humidity range. Furthermore, the effect of temperature, although it can be considered through concrete resistivity to some extent, is ignored. It should be noted that temperature directly affects the kinetic parameters of steel corrosion, and for accurate prediction of corrosion rates, the effect of temperature variations should be considered in the predictive models.

4.2 *Mathematical models*

Mathematical models are based on either the mathematical representation of controlling mechanisms of corrosion or numerical modelling. In the following sections, prominent mathematical models from both categories are presented without providing details.

4.2.1 *Raupach and Gulikers (1999)*

Raupach and Gulikers (1999) used an equivalent electric circuit approach to model steel corrosion in concrete. In this approach, the macro-cell current flowing from anodic to cathodic sites is calculated using Ohm's law by dividing the potential difference between passive and active sites on the steel surface to the equivalent circuit resistance, which is the sum of concrete resistance, and anodic and cathodic

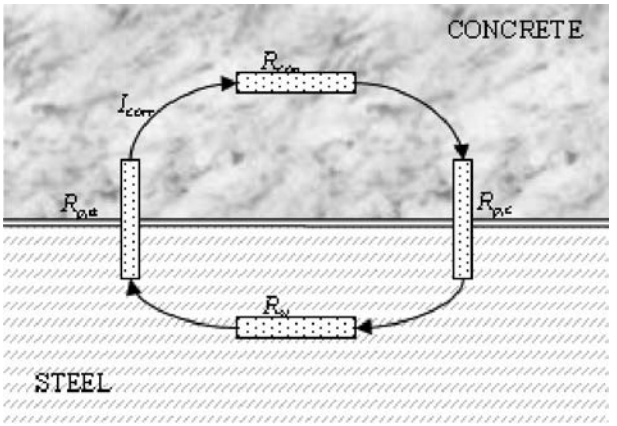


Figure 2 Illustration of the equivalent circuit used by Raupach and Gulikers (1999).

polarization resistances. A schematic of a typical equivalent circuit is illustrated in Fig. 2.

Using the equivalent circuit approach, the following model was proposed:

$$I_{corr} = \frac{E^{\circ}_{O_2} - E^{\circ}_{Fe}}{R_{p,c} + R_{con} + R_{p,a} + R_{st}} \tag{21}$$

where E°_{Fe} is the anodic equilibrium potential, $E^{\circ}_{O_2}$ is the cathodic equilibrium potential, $R_{p,c}$ is the cathodic polarization resistance, $R_{p,a}$ is the anodic polarization resistance and R_{st} is the electrical resistance of steel. In Eq. 21 resistance of steel can be neglected due to high conductivity, and the polarization resistance of anode and cathode can be calculated using Butler-Volmer kinetics (Brad and Faulkner 2001).

Although equivalent circuit model considers activation and resistance control mechanisms of corrosion, its application to real-life situations is very limited because application of the model requires a priori knowledge the exact distribution of the anodic and cathodic sites on the steel surface. In addition, this model is not practical for the prediction of corrosion rates of reinforced concrete members that contain complex reinforcement detailing and geometry.

4.2.2 *Isgor and Razaqpur (2001)*

Isgor and Razaqpur (Isgor, 2001; Isgor and Razaqpur, 2006) used a nonlinear finite element solution algorithm to calculate the rate of corrosion (current density) on the steel surface by solving the governing differential equation of the electrical potential distribution of the corroding system. The equation governing the potential distribution can be derived from first principles (Munn, 1991). Assuming electrical charge conservation, the potential distribution can be represented by the Laplace's equation:

$$\nabla^2 \phi = 0 \tag{22}$$

Calculation of the potential distribution on the surface of the steel involves the solution of Eq. 22, subject to prescribed boundary conditions. These boundary conditions comprise the relationship between potential and current density for the anodic and cathodic regions as well as prescribed current densities. For the anodic and cathodic regions of the steel (see Fig. 3 for an example domain), the boundary condition are defined as $\phi = \phi_a$ and $\phi = \phi_c$, where ϕ_a and ϕ_c are polarized anodic and cathodic potentials which can be expressed as follows (Stern and Geary, 1957):

$$\phi_a = \phi_{Fe}^{O} + \beta_a \log \frac{i_a}{i_{oa}} \quad (\text{on } \Gamma_a) \tag{23}$$

$$\phi_c = \phi_{O_2}^{O} + \beta_c \log \frac{i_c}{i_{oc}} - \frac{2.303RT}{zF} \log \frac{i_L}{i_L - i_c} \quad (\text{on } \Gamma_c) \tag{24}$$

where ϕ_{Fe}^{o} and $\phi_{O_2}^{o}$ are the standard half-cell potentials of Fe and O_2, respectively, β_a is the Tafel slope of the anodic reaction, and i_{oa} is the anodic exchange current density, β_c is the Tafel slope of the cathodic reaction, i_{oc} is the exchange current density of the cathodic reaction, and i_L is the limiting current density. Equations 23 and 24, respectively, represent the polarization behaviour of the anodic and cathodic sites that form on the steel surface. In the model, Isgor and Razaqpur assumed that anodic sites polarize due to activation polarization, and cathodic sites due to activation and concentration polarization.

Because of the non-linearity in the boundary condition, they used a non-linear solution algorithm to obtain the potential distribution in the domain of interest. Once the potential distribution is accurately determined, the corrosion rate distribution (i.e. current density) along the steel surface can be calculated using

$$i = -\frac{1}{\rho} \frac{\partial \phi}{\partial n} \tag{25}$$

where n = direction normal to the rebar surface. Further details on this model can be obtained from Isgor and Razaqpur (2006).

There are two main limitations of Isgor and Razaqpur's model: Firstly, the numerical solution of the nonlinear problem can be quite expensive, especially for complicated geometries and under conditions which lead to significant nonlinearity in the boundary conditions. Some conditions that can lead to significant nonlinearity can be listed as: (1) extremely low or high values of concrete resistivity, (2) presence of large concentration polarization (as in the case of saturated concrete in which the oxygen diffusion is very slow), and (3) very small anode-to-cathode area ratios. Secondly, the prediction of the anodic and cathodic locations on the steel surface can be quite difficult and may require the coupled solution of initiation and propagation stage processes simultaneously. Because of these two limitations, although it is very suitable for research studies, Isgor and Razaqpur's model might be considered as difficult to use for the practicing engineers who do not have expertise in the area of corrosion of steel in concrete.

4.2.3 *Gulikers (2005)*

Using a similar electric circuit approach that was previously developed by Raupach and Gulikers (1999) for calculating the potential difference between anodic and cathodic

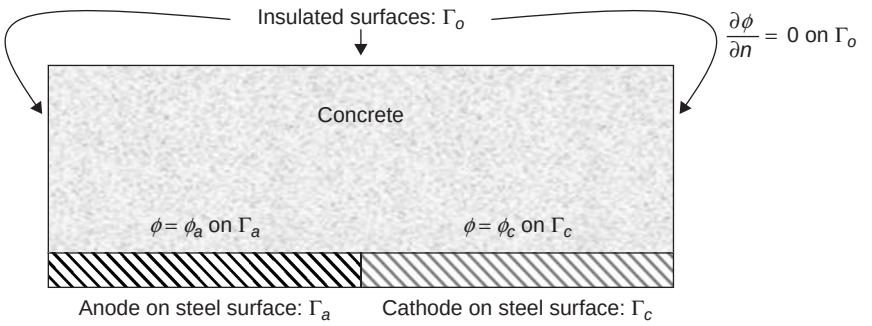


Figure 3 Boundary conditions of a defined as Dirichlet type in Isgor and Razaqpur model (Isgor, 2001).

sites, Gulikers (2005) developed a relationship between corrosion current density and concrete resistance. Concrete resistance then was converted to concrete resistivity by considering a geometry (i.e. shape) factor. This geometry factor was found by comparing the equations resulting from equivalent circuit model and the model proposed by Alonso et al. (1988); the value for the geometry factor suggested was 578×10^{-3}. The model is given by:

$$i_{corr} = \frac{F_G^{-0.8125} 98.696 \times 10^{-3}}{\rho^{0.8125}} \tag{26}$$

where i_{corr} is in (A/m^2) and F_G (m^{-1}) is the geometry factor.

Gulikers' model shares the same disadvantages of other similar models as it fails to predict the corrosion behaviour under diffusion control. Furthermore, Gulikers's model is constructed considering constant kinetic parameters of concrete which may vary substantially due to environmental effects such as temperature, relative humidity and concrete properties such as w/c and porosity. Therefore, its applicability can be quite limited in real-life applications.

4.2.4 *Pour-Ghaz and Isgor (2007)*

Pour-Ghaz and Isgor (Pour-Ghaz, 2007) improved the nonlinear solution algorithm developed by Isgor and Razaqpur (2006). The new version has been demonstrated to solve cases with significant non-linearity (e.g. due to significant concentration polarization or low resistivity) and various geometric and boundary conditions, for which converged solutions were not previously possible. The improved solution algorithm allowed Pour-Ghaz and Isgor to use the solution algorithm as a virtual experiment tool to analyze thousands of cases. They stated that a comprehensive experimental study to investigate the effects of all the necessary parameters on the corrosion of steel in concrete and to develop a comprehensive closed-form regression model simulating the existing corrosion rate measurement techniques (such as the ones that are based on polarization resistance method) is not practically feasible.

Due to the large number of variables affecting the phenomenon, it was stated that a large number of tests (in hundreds) need to be carried out to obtain data that will provide a regression-based model that is statistically reliable. Considering the fact that there are different available devices that are based on the same theoretical basis but measure corrosion rates with relatively large differences, to include the uncertainty associated with the device used, most of the tests need to be repeated using several devices.

Using the results of the thousands of virtual experiments, Pour-Ghaz and Isgor developed two models for predicting steel corrosion in concrete. The first model (Pour-Ghaz, 2007), which is presented in Eq. 27, uses half-cell test data along with a number of environmental parameters as input to predict the corrosion rate:

$$i_{corr} = \left(\frac{17.043 \times 10^3}{\phi_{ave}^2}\right)\left(-9.124 \times 10^{-8} d + 1.718 \times 10^{-7}\frac{T}{\rho} + 1.576 \times 10^{-7}\Delta\phi_{surf}\right) - 2.043 \times 10^{-3} \tag{27}$$

where the terms, ϕ_{ave} [volts] and $\Delta\phi_{surf}$ [volts], can be obtained from regular half-cell tests; the other terms, ρ[Ω-m], T[°K] and d[m] can be easily determined using non-destructive methods. Although ϕ_{ave} is what current half-cell tests provide, $\Delta\phi_{surf}$ is an additional reading that needs to be collected from the test. These investigators defined $\Delta\phi_{surf}$ as the absolute value of the maximum potential difference that is obtained from a number of half-cell measurements carried out within a 300 mm radius from a point on the steel reinforcement. As a guideline, it was suggested that the measurements be carried out at six equally-spaced locations (separated by 50 mm) along the reinforcement. However, more accurate readings of $\Delta\phi_{surf}$ can be obtained by increasing the number of test points, especially when pitting corrosion on the steel is suspected.

The second model, which does not depend on half-cell measurements, includes the effects of temperature, oxygen availability (through limiting current density), cover thickness and concrete resistivity, and allows the determination of average corrosion rate (that is normalized over the anodic surface of the steel) and maximum corrosion rate (that corresponds to the maximum corrosion rate on the steel surface). The model is given as follows:

$$\left\langle \begin{matrix} i_{corr,ave} \\ i_{corr,\max} \end{matrix} \right\rangle = \frac{1}{\tau\rho^{\gamma}} \begin{pmatrix} \eta T d^{\kappa} i_L^{\lambda} + \mu T \nu^{i_L^{\varpi}} \\ +\theta(T i_L)^{\vartheta} + \chi\rho^{\gamma} + \zeta \end{pmatrix} \tag{28}$$

where a description and suggested values of the various coefficients of this equation (i.e. $\tau, \gamma, \eta, \kappa, \lambda, \mu, \nu, \varpi, \theta, \vartheta, \chi, \zeta$) can be found in Pour-Ghaz et al. (2008).

Although Pour-Ghaz and Isgor model has a wide range of applicability, the simulations that provided the data for the regression model given in Eq. 28 have been carried out using constant corrosion parameters (e.g. Tafel slopes, exchange current densities, etc.) Therefore, if different corrosion parameters are to be used, the simulations, and hence the regression analysis, need to be carried out again to obtain a revised closed-form equation. This can be considered as a limitation of the Pour-Ghaz and Isgor model.

Figure 4 Room 1 (Window is along the top axis of the plot).

5 A case study

This case study deals with the prediction of corrosion rate in an apartment building and is intended to demonstrate the combined field measurement and analytical methods. The concrete in parts of the building is heavily contaminated with calcium chloride used as accelerator in the fresh concrete during construction. Free chloride content as high as 1.4% by weight of cement was detected. The building envelope is believed to inadequately shield the building from external moisture infiltration at some locations. The infiltrated moisture, together with moisture from internal sources, is absorbed by concrete and raises its moisture content sufficiently to cause steel reinforcement corrosion in the some parts of the floor slab. The corrosion has led to cracking, spalling and delamination of concrete in certain parts of the building.

Based on available information and the observations made on site, the current case study makes the following assumptions:

- The concrete is pre-contaminated, with many areas of the floor slab having chloride content exceeding 0.9%, and this contamination has resulted in the depassivation of the steel. The delamination and spalling experienced in certain units are due to the expansive stresses created by the corrosion of steel in concrete.
- Other deteriorative mechanisms may be also be responsible for the observed damage, but have not been considered in this study.

5.1 *Field measurements*

For two sites in the building, the following quantities were measured in Spring 2007: (1) Corrosion rate by Galvanostatic pulse technique; (2) Half-cell potential using Ag/AgCl half cell; and (3) Resistivity of concrete. Figs. 4 and 5 show the floor slab of the two rooms in the building that were investigated. In Room 1, a test grid of 81 points (300 mm $\times$ 300 mm grid division) while in Room 2 a 16 point (600 mm $\times$ 600 mm grid

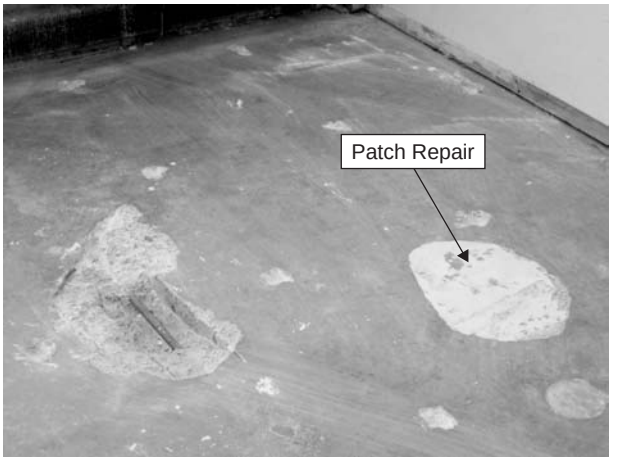


Figure 5 Room 2 and the patch repair (Window is along the top axis of the plot).

division) grid was used. A relatively large patch repair in Room 2 was observed (See Figures d 5). The measured half cell potential, resistivity and corrosion rates for the two rooms are plotted in Figs. 6 and 7 for Rooms 1 and 2, respectively.

It should be noted that the results presented in Figs. 6 and 7 are valid at the time of the testing. Half-cell, galvanostatic pulse and resistivity measurements are very sensitive to temporal environmental conditions; hence conclusive interpretations of the test data can only be made when the site is tested over a long period of time that covers seasonal as well as daily variations in temperature and humidity. Consequently, the collected data and observations should be interpreted accordingly.

Fig. 6(a) shows the half-cell potential distribution in Room 1. Based on the ASTM Standard C876 these readings can be used to assess the probability of corrosion in the particular region of the floor slab, as described previously in this paper. The potential values in Fig. 6(a) are given with respect to Ag/AgCl standard electrode, which can be converted to values with respect to copper/copper sulfate standard electrode (CSE) by subtracting 117 mV from the indicated values in the figure. Thus the half cell potential values in Fig. 6(a) with respect to the CSE would range from −157 mV to −477 mV, which clearly indicates that there is greater than 90% probability of corrosion over a large portion of the floor. In other words, any area with a potential less than −233 mV in Fig. 6(a) has 90% probability of corrosion. It is noteworthy that the regions with most probable corrosion activity are located along one of the walls. Note that the potential values do not provide any information about the rate of corrosion; they simply indicate whether corrosion is likely to occur.

The concrete resistivity values for the floor slab in Room 1 are shown in Fig. 6(b). These values are quite high, i.e. greater than 1000 Ω-m in most parts of the floor. Accordingly, the corrosion rate measured using the galvanostatic pulse technique generally yielded low values in a range from 0.06–0.36 $\mu A/cm^2$. The higher corrosion rates (e.g. ∼0.36 $\mu A/cm^2$) were observed in zones where the concrete had previously experienced extensive cracking and spalling. It was also observed that the corrosion

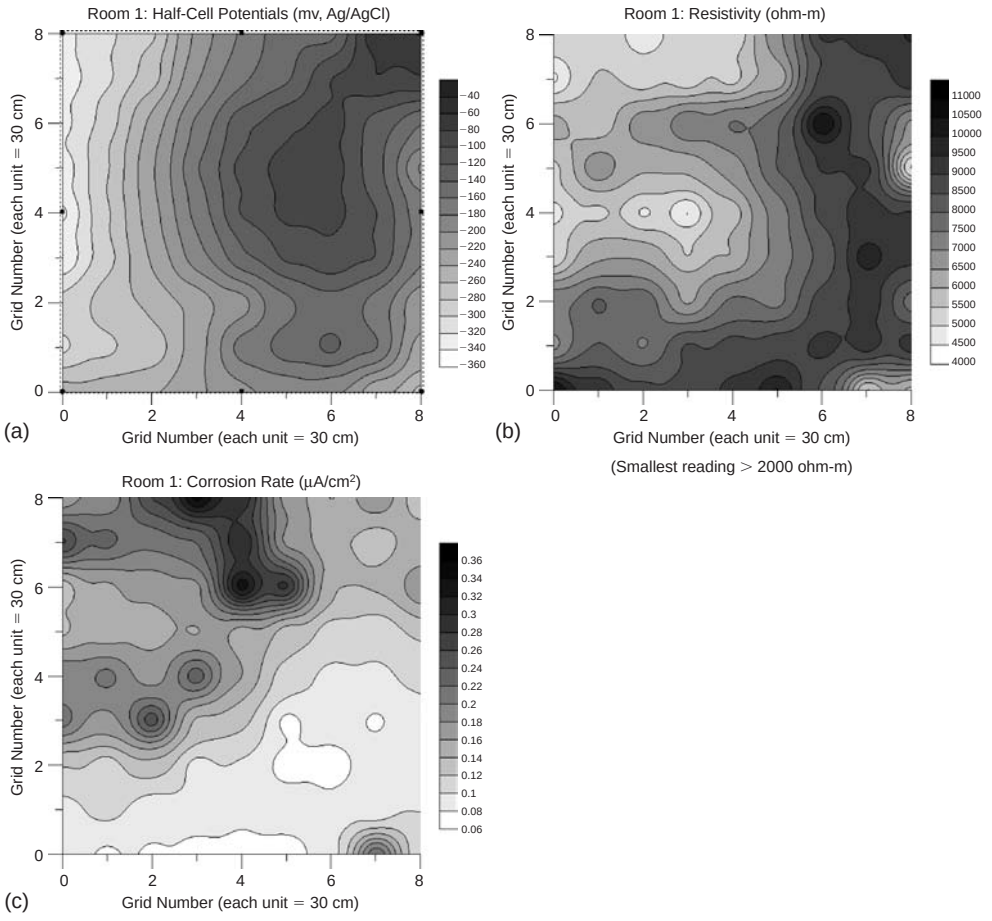


Figure 6 Field measurements in Room 1 (a) Half-cell, (b) resistivity, (c) corrosion rate (Window is along the top axis of the plot).

rates are generally higher closer to the window. The high resistivity and relatively low corrosion rates observed in Room 1 at the time of testing can be attributed to the fact that the tests were carried out in spring, just after a winter with very low relative humidity. It should also be noted that prior to taking measurements, the ambient air temperature outside the building was for days only in the single digit C°. In addition, the apartment was unoccupied and therefore there was less likelihood of water leakage from the kitchen and the bathroom.

The measurements in Room 2, on the other hand, revealed an interesting behaviour. Although the resistivity of most of the floor slab in this room was of the same order of magnitude as in Room 1, there were some locations with very low resistivity (e.g., 180 Ω-m). It is interesting to note that the zones with low resistivity are the same zones where the repair patch is located. As it can be observed from Fig. 7(c), the corrosion rate of the rebars in the vicinity of repair patch is significantly larger than the other parts of the room, reaching up to 6.5 $\mu A/cm^2$, which translates into a bar diameter

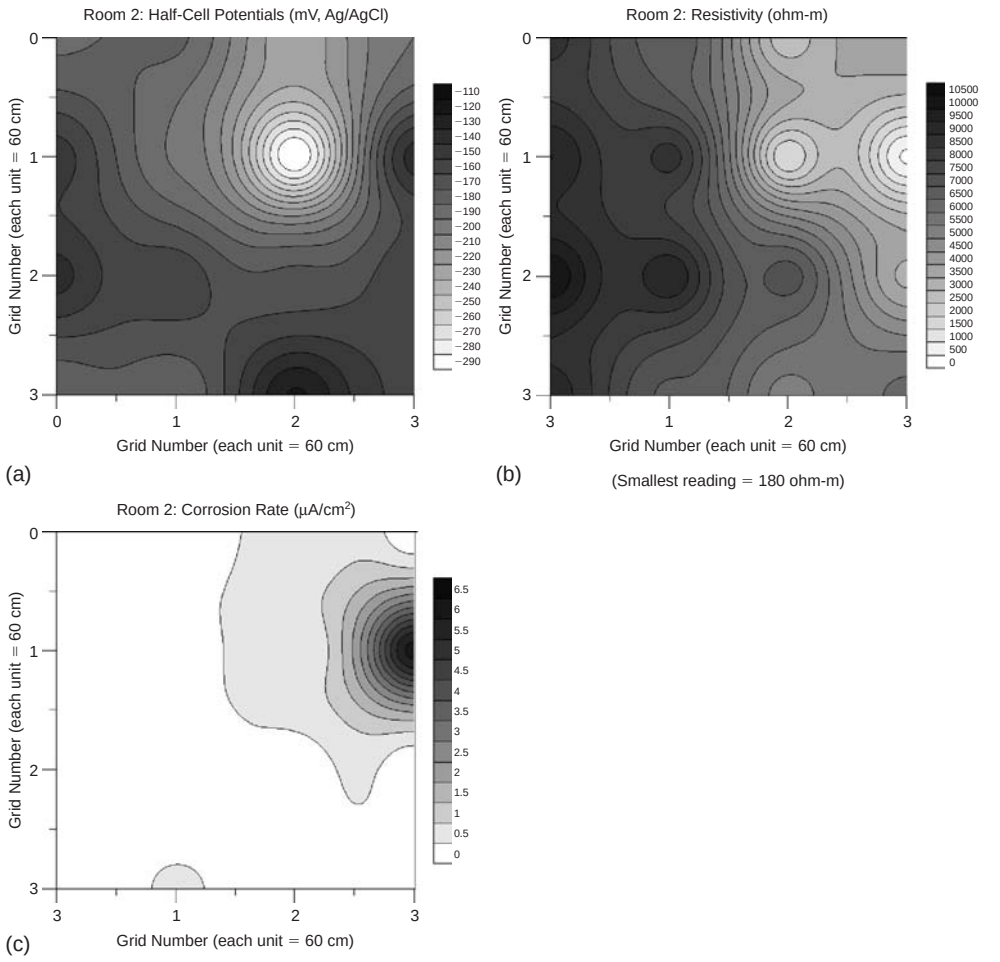


Figure 7 Field measurements in Room 2 (a) Half-cell, (b) resistivity, (c) corrosion rate (Window is along the top axis of the plot).

loss of 78 μm/year. It should be stated that 70 μm loss of diameter is adequate to initiate cracking of concrete cover. Fig. 7(c) illustrates that there is significant macro-cell activity in the room, where a small part of the reinforcement grid acts as an active anode and the rest of the mesh acts as a cathode. This observation in Room 2 is important in the sense that it shows, even after a winter with low relative humidity, there may be localized sections of concrete with low resistivity, facilitating the formation of local macro-cells with relatively large corrosion rates.

5.2 *Corrosion rate prediction*

The corrosion rates of the critical areas of the slabs in the two rooms, which showed large corrosion rates, have been determined using two empirical (i.e., Morinaga, 1990;

Table 1 Corrosion parameters used in the mathematical models.

Standard cathode potential (ϕ^0_{O2})	0.160V (SCE)
Standard anode potential (ϕ^0_{Fe})	−0.780V (SCE)
Cathodic Tafel slope (β_c)	−0.180V/dec
Anodic Tefel slope (β_a)	0.0895V/dec
Cathodic exchange current density (i_{oc})	10×10^{-6} A/m^2
Anodic exchange current density (i_{oa})	300×10^{-6} A/m^2

and Liu and Weyers, 1998) and three mathematical (i.e., Isgor and Razaqpur, 2001; Gulikers, 2005; and Pour-Ghaz and Isgor, 2007) predictive models described in Section 4. For Room 1, the area near the window with high corrosion activity (see Fig. 6(c)) is modeled. For Room 2, since the most significant corrosion activity was observed in the patch repair (see Fig. 7(c)), this area was simulated using the predictive models. The resistivity readings for both rooms were obtained from Figs. 6(b) and 7(b); the resistivities of the modeled areas in Rooms 1 and 2 were 4000 Ω-m and 50 Ω-m, respectively. It is assumed that both rooms are kept at 25°C. For some of the empirical models, the following additional assumptions were made: (1) w/c = 0.40; (2) concrete cover thickness (d) = 25 mm; (3) diameter of steel (d_{st}) = 20 mm; age of concrete = 35 years; chloride contamination = 1% by weight of cement; and oxygen concentration around the steel (C_{O2}) = 0.00125 kg/m^3. The corrosion parameters that are used in the Isgor and Razaqpur model are reported in Table 1.

The comparison of corrosion rate predictions from empirical and mathematical models with experimental data for Rooms 1 and 2 are presented in Figs. 8 and 9, respectively. By observing the difference in the range of corrosion rates between the two rooms, it can be stated that corrosion activity in Room 2 is significantly (almost by two orders of magnitude) higher than that in Room 1. This is mainly due to the difference between the resistivities of the rooms. The patch repair in Room 2 had a very low resistivity, leading to high corrosion rates.

As illustrated in Fig. 8, the corrosion rate predictions of the selected empirical and mathematical models in Room 1 are in the same order of magnitude with the experimental measurement. Except for Gulikers' model (2005), which underestimates the corrosion rate by half, the predictions of the models are quite accurate for Room 1. It should be noted that the resistivity of Room 1 was 4000 Ω-m, which is within the range of applicability of the empirical models used in this comparison.

On the other hand, as illustrated in Fig. 9, the corrosion rate predictions by the selected empirical models for Room 2 significantly underestimate the actual corrosion rates. This is mainly due to the fact that these empirical models were obtained from experiments in which concrete resistivity was significantly larger than the resistivity of the patch repair. Therefore, at low resistivities, the empirical models fail to predict the corrosion rates accurately. Actually, by comparing Figs. 8 and 9, it can be observed that these empirical models do not consider the effect of concrete resistivity as a major factor affecting corrosion rates; therefore, they predict almost the same corrosion rates for both rooms.

However, mathematical models, particularly those developed by Isgor and Razaqpur (2001) and Pour-Ghaz and Isgor (2007), predict the corrosion rates quite accurately.

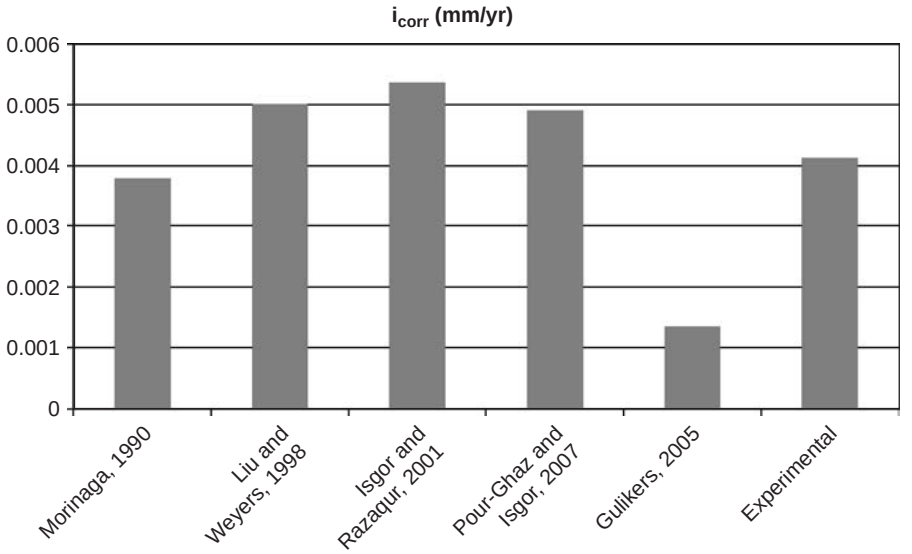


Figure 8 Comparison of corrosion rate predictions from empirical and mathematical models with experimental data for Room 1.

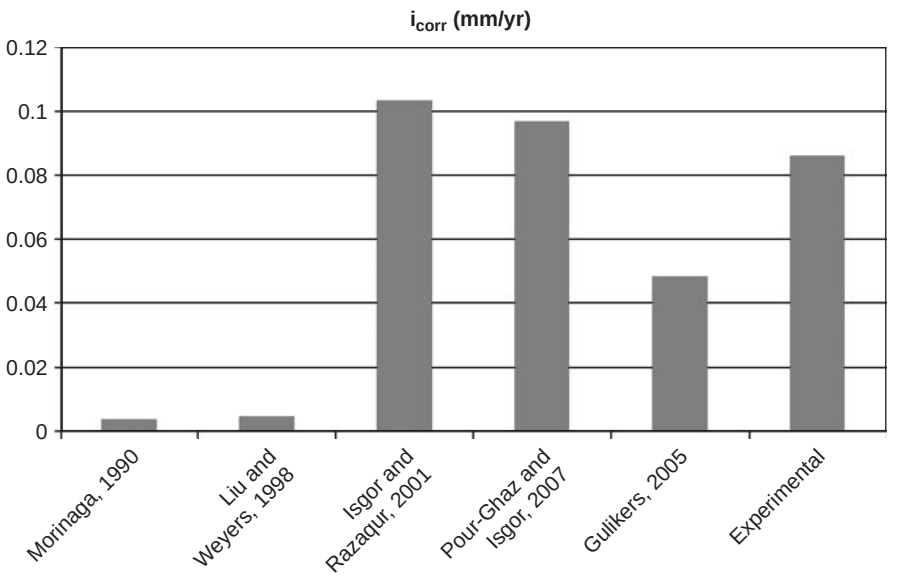


Figure 9 Comparison of corrosion rate predictions from empirical and mathematical models with experimental data for Room 2.

Note that Gulikers' model (2005) underestimates the corrosion rate for Room 2 as it did for Room 1.

It is important to state, based on these results, that it is possible to predict corrosion rates using empirical and mathematical models. However, each model has deficiencies, and the interpretation of the predictions must be made carefully by considering these limitations. In addition, for reliable corrosion rate predictions, it is important to use accurate input data (e.g. concrete resistivity, temperature, chloride content); therefore it is always necessary to complement predictive models with accurate experimental measurements. Simulations using predictive models, combined with relatively infrequent corrosion rate measurements, can be the answer to health monitoring of reinforced concrete infrastructure.

6 Conclusions

In order to accurately monitor the corrosion rates in concrete, the corrosion rate needs to be measured frequently for a relatively long period of time so that temporal variations can be captured. Therefore, although existing corrosion measurement techniques are quite practical for instantaneous monitoring of structures, they become significantly expensive and impractical if the monitoring is to be done over long periods. In addition, the common corrosion measurement techniques have inherent deficiencies, which make it difficult to derive reliable engineering conclusions that are solely based on their output.

Modeling of steel corrosion in concrete structures is an important tool that can help in the better interpretation of the data from corrosion measurement techniques. Although it is difficult, if not impossible, to completely replace the nondestructive testing methods with theoretical models, accurate simulations combined with relatively infrequent, and hence inexpensive, corrosion rate measurements can be the answer to health monitoring of reinforced concrete infrastructure. It is contented in this paper that a combination of field measurements and numerical models may be the best way to obtain a reasonable estimate of corrosion activity in a structure over the long-term.

Acknowledgements

The authors would like to thank Halsall Associates Limited of Ottawa for the opportunity to carry out corrosion measurement tests in the apartment building complex and to Mohammad Pour-Ghaz and Pouria Ghods, who helped during the field testing of the case study.

References

Ahmad, S., Reinforcement Corrosion in Concrete Structures, its Monitoring and Service Life Prediction – A Review, *J . of Cement and Concrete Composites*, No. 25, Elsevier Press, pp. 459–471, 2003.

ASTM C876-91, Standard Test Method for Half-Cell Potential of Reinforcing Steel in Concrete, *American Society for Testing Materials*, 03(02), pp. 434–439, 1991.

Bazant, Z.P., Physical Model for Steel Corrosion in Concrete Sea Structures – Application, *ASCE J. Struct. Div.*, 105(6), 1155–1166, 1979.

Berkeley, K.G.C., and Pathmanaban, S., *Cathodic Protection of Reinforcement Steel in Concrete*, Butterworth-Heinemann, London, 1990.

Bertolini, L., Elsener, B., Pedeferri, P., and Polder, R., *Corrosion of Steel in Concrete*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2004.

Brad, A.J., and Faulkner, L.R. *Electrochemical Methods Fundamentals and Applications*, 2nd Ed., John Wiley & Sons, New York, 2001.

Broomfield, J.P., *Corrosion of Steel in Concrete*, E&FN Spon, London, 1997.

DuraCrete, *Modelling of Degradation*, Document BE95-1347/R4-5, Gouda, 2000.

Elsener, B., Macrocell Corrosion of Steel in Concrete – Implications for Corrosion Monitoring, *J. of Cem. Concr. Compos.*, 24(1), 65–72, 2002.

Elsener, B., Corrosion Rate of Steel in Concrete – Measurements beyond the Tafel Law, *Corr. Sci.*, 47(12), 3019–3033, 2005.

Ge, J., On the Numerical Solution of Laplace's Equation with Nonlinear Boundary Condition for Corrosion of Steel in Concrete." *M.A.Sc. Thesis*, Carleton University, Ottawa, Canada, 2006.

Ghods, P., Isgor, O.B., and Pour-Ghaz, M. A Practical Method for Calculating the Corrosion Rate of Uniformly Depassivated Reinforcing Bars in Concrete, *Mater. Corros.*, 58(4), 265–272, 2007.

Gulikers, J., Theoretical Considerations on the Supposed Linear Relationship between Concrete Resistivity and Corrosion Rate of Steel Reinforcement, *Mater. Corros.*, 56(6), 393–403, 2005.

Hunkeler, F., Corrosion in Reinforced Concrete: Processes and Mechanisms, Chapter 1, *Corrosion in Reinforced Concrete*, Edited by Hans Böhni, CRC Press, Washington DC, 2005.

Isgor, O.B., A Durability Model for Chloride and Carbonation Induced Steel Corrosion in Reinforced Concrete Members, *Ph.D. Thesis*, Carleton University, Ottawa, Canada, 2001

Isgor, O.B., and Razaqpur, A.G., Modelling Steel Corrosion in Concrete Structures." *Mater. Struct.*, 39(287), 291–302, 2006.

Liu, T., and Weyers, R.W., Modelling the Corrosion Process in Chloride Contaminated Concrete Structures." *Cem. Concr. Res.*, 28(3), 365–379, 1998.

Luping, T., *Calibration of the Electrochemical Methods for the Corrosion Rate Measurement of Steel in Concrete*, NORDTEST Project No. 1531-01, SP REPORT 2002:25, Swedish National Testing and Research Inst., 2002.

Morinaga, S., Prediction of Service Life of Reinforced Concrete Buildings Based on the Corrosion Rate of Reinforcing Steel." *Proc. of the 5th International Conference on Durability of Building Materials and Components*, Brighton, U.K., 1990.

Munn, R.S., and Devereux, O.F., Numerical Modeling and Solution of Galvanic Corrosion Systems. Part I., *Corrosion*, 47(8), 612–618, 1991.

Neville, A.M., *Properties of Concrete*, 4th Ed., Pearson Education Ltd., NY, 1995.

RILEM TC 154-EMC, Half-cell Potential Measurements – Potential Mapping on Reinforced Concrete Structures, *Mater. Struct.*, 36(7), 461–471, 2003.

Pour-Ghaz, M. A Novel Approach for Practical Modelling of Steel Corrosion in Concrete, *M.A.Sc Thesis*, Carleton University, Ottawa, Canada, 2007.

Raupach, M., and Gulikers, J., A Simplified Method to Estimate Corrosion Rates – A New Approach Based on Investigations of Macrocells, *8th International Conference on Durability of Building Materials and Components*, Vancouver, Canada, 1999.

Stern, M., and Geary, A.L., Electrochemical Polarization, I. A Theoretical Analysis of the Shape of Polarization Curves, *J. Electrochem. Soc.*, 104(12), 56–63, 1957.

Uhlig, H.H., and Revie, R.W., *Corrosion and Corrosion Control*, 3rd Ed., John Wiley & Sons, New York, 1985.

Chapter 4

Introduction of optimum structural design and its recent developments

Franklin Y. Cheng

Missouri University of Science & Technology, (formerly University of Missouri – Rolla), Rolla, MO, USA

ABSTRACT: During the last four decades a significant amount of research has been published in the area of structural optimization. These papers are dealing with development of algorithms and their advantages in numerical solutions along with engineering applications. Chronologically the publications may be roughly classified as follows: mathematical programming, optimality criteria, and genetic algorithms. The objective functions (cost function) are also comprehensively investigated and may also be classified as: minimum weight, minimum cost, and multi-objective function. This paper introduces the fundamentals related to the algorithms, cost functions, constraints, and numerical procedures of the aforementioned optimization techniques and then advances to sophisticated application of the techniques for optimum design of various structural systems. It is hoped that the practicing engineers can closely follow the developments of various methods and appropriately select them as the powerful design tool for their needs in order to achieve better design in the real construction world.

1 Introduction

It has long been recognized that conventional structural designs are based on repeated analyses with assumed stiffness of the constituent members of a given structure. If the preliminary stiffnesses are not correctly ascertained, a poor design will be the result in spite of the number of analysis cycles and the sophistication of the analysis computer programs. Apparently, conventional design is an art of which the processes cannot guarantee an efficient design. Therefore, if a design is to be reliable, it should be based on the logical mathematical derivation of optimum design procedures from which an economic and serviceable structure can be obtained. Thus the results of an optimum design should satisfy a set of constraints, such as stresses, displacements, frequencies, buckling loads, member sizes, and dynamic forces, as well as the story drifts. In other words, the optimization procedure can redistribute the stiffness of a system according to the constraint and loading requirements. In the past a significant amount of research has been done in the area of structural optimization. This paper briefly introduces most of the optimization algorithms currently in vogue and their applications.

2 Description of nonlinear optimization

2.1 *Constraint vs. unconstraint and nonlinear vs. linear problems*

The typical mathematical optimization problem can be stated as follows:

$$\begin{aligned} &\text{Minimize} && F(x) && (1)\\ &\text{Subject to} && g_j(x) \geq 0 \quad j = 1, \ldots, m && (2)\\ &&& \overline{x}_i \geq x_i \geq \underline{x}_i \quad i = 1, \ldots, n && (3)\\ &\text{Find} && \mathrm{X}^* = [x_1^*, \ldots, x_n^*]^{\mathrm{T}} && (4) \end{aligned}$$

where x_i = design variables such as cross-sectional sizes (area, moment of inertia, etc.) of structural members; $F(x)$ = objective function such as structural weights, construction cost, etc.; $g_j(x)$ = constraints representing structural response limitations (stresses, displacements, frequencies, buckling load, etc.); $\underline{x}_i$ and $\overline{x}_i$ = reasonable member sizes, analysis validity consideration, etc. m = number of constraints; and n = number of design variables. Eqs. (1) through (4) represent a class of constrained optimization problem for which the solution procedures may be expressed in Eqs. (5) through (8): Lagrangian function, Arora, 2004.

$$\left.\begin{aligned} L(x, \lambda) &= F(x) - \sum_{j=1}^{m} \lambda_j g_j \\ \lambda_j &\geq 0 \quad j = 1, \ldots, m \end{aligned}\right\} \tag{5}$$

Kuhn–Tucker condition (KT):

$$\frac{\partial F}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \quad \text{if } \underline{x}_i < x_i < \overline{x}_i \tag{6}$$

$$\frac{\partial F}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} > 0 \quad \text{if } x_i = \underline{x}_i \tag{7}$$

$$\frac{\partial F}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} < 0 \quad \text{if } x_i = \overline{x}_i \tag{8}$$

When the objective functions and constraints are linear, then the statement problem expressed in Eqs. (1) through (4) is called a linear programming problem. If any of the problem function is nonlinear, then the statement is a nonlinear programming problem. Solutions to nonlinear programming problem are problem dependent and are much less standard than that to linear programming (Cheng, Venkayya, and Khachaturian, 1976).

If Eq. (2) is not considered, then we have

$$\left.\begin{aligned} &\text{Minimize} && F(x) \\ &\text{Subject to} && \overline{x}_i \geq x_i \geq \underline{x}_i \end{aligned}\right\} \tag{9}$$

which is unconstrained minimization problem. Many constrained problems are solved by solving a sequence of unconstrained problems. The problem in Eqs. (1) through (4) is called convex programming problem when the functions $\{F(x); -g_j(x); j = 1, \ldots, m\}$ are convex. Then the feasible region is a convex set for which the local solution is also global.

2.2 *Primal vs. dual problems*

Define primal points satisfy the side constraints

$$X = \{\mathbf{x} : \underline{x}_i \leq x_i \leq \overline{x}_i; \quad i = 1, 2, \ldots, n\} \tag{10}$$

and define set of dual points satisfying the non-negativity conditions as

$$\Lambda = \{\boldsymbol{\Lambda} : \lambda_j \geq 0; \quad j = 1, 2, \ldots, m\} \tag{11}$$

The primal problem is expressed as

$$\text{Minimize} \quad F(x) \quad \text{for} \quad x \in \mathrm{X} \tag{12}$$

$$\text{Subject to} \quad g_j(x) \geq 0; \quad j = 1, 2, \ldots, m \tag{13}$$

When $F(x)$ is strictly convex and $g_j(x)$ are concave, there exists a unique dual problem. The solution of the dual problem can be obtained through a two phase procedure as follows:

$$\begin{array}{ll} \max & \min L(x, \lambda) \\ x \in X & \lambda \in \Lambda \end{array} \tag{14}$$

where

$$L(x, \lambda) = F(x) - \sum_{j=1}^{m} \lambda_j g_j(x) \tag{15}$$

The dual problem can be written as

$$\left.\begin{array}{ll} \text{Maximize} & l(\lambda) \\ \text{Subject to} & \lambda_j \geq 0 \quad j = 1, 2, \ldots, m \end{array}\right\} \tag{16}$$

where

$$l(\lambda) = \min_{x \in X} \mathrm{L}(x, \lambda) \tag{17}$$

In primal problem there are n variables, m general constraints and $2n$ side constraints. While in dual problem there are only m dual variables and m dual non-negativity

constraints. The relationships may be summarized as

	Primal problem	Dual problem
Variables	n	m
General Constraints	m	m
Side Constraints	$2n$	

which implies that the dual problem is easier to solve than the prime. This behavior is similar to that of linear optimization which will be illustrated later (see sub-section 4.1.2).

3 Basis of nonlinear minimization algorithms

Most algorithms for solving minimization problem are numerical methods and are iterative because the analytical methods for solving such problems are difficult and sometimes impossible. In fact most of the unconstrained minimization methods solved by sequel approach are decent methods of numerical procedures expressed as follows

$$F(\mathbf{x}^{(k+1)}) < F(\mathbf{x}^{(k)}) \tag{18}$$

where k is the iteration number. It is apparent that minimizing $F(\mathbf{x})$ is to successively to find the search directions $\boldsymbol{s}^{(k)}$ so that

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\boldsymbol{s}^{(k)} \tag{19}$$

$\boldsymbol{s}^{(k)}$ must be a downhill direction for a sufficiently small $\alpha > 0$. Thus Eq. (18) should hold as

$$F(\mathbf{x}^{(k)} + \alpha\mathbf{s}^{(k)}) < F(\mathbf{x}^{(k)}) \tag{20}$$

An equivalent requirement of Eq. (20) is that

$$\boldsymbol{s}^{(k)\mathrm{T}}\nabla\boldsymbol{F}(x^{(k)}) < 0 \tag{21}$$

$\nabla F(\mathbf{x}^{(k)})$ is the gradient of $F(\mathbf{x})$ at $\mathbf{x}^{(k)}$. This is the basic descent algorithm. The process of finding $\alpha^{(k)}$ by estimating a minimum of a function, $\phi(\alpha)$, is called line search. Various techniques are developed for this purpose.

The method called steepest descent based descent algorithms is the fundamental method in nonlinear programming. It provides the basis for all gradient methods. In this method the downhill directions are taken as the negative gradient vectors.

$$\left.\begin{aligned} &s^{(k)} = -g^{(k)} \\ &s^{(k)\mathrm{T}}g^{(k)} < 0 \quad \text{if } g^{(k)} \neq 0 \end{aligned}\right\} \tag{22}$$

So, from the point $x^{(k)}$, we search along the direction of negative gradient to a minimum point $x^{(k+1)}$ on this line. It converges fast for nearly circular contours. If the contours are very elongated the convergence will be slow. A significant amount of literature is available for improvement of search techniques such as conjugate

gradient method, Newton's method, Davidon-Fletcher-Powell method as well as Broyden-Fletcher-Goldfarb-Shanno method (Cheng 1986).

4 Optimization methods for constrained problems and sample illustrations

There are many practical methods to solve the constrained nonlinear programming problems including a) linearization method, b) primal method, and c) transformation method. The most natural approach and easy understanding of a nonlinear problem is to replace it with a sequence of linear programming problems. In this section, these three methods are respectively introduced in sub-sections 4.1, 4.2, and 4.3, with sample numerical illustrations.

4.1 *Linear programming for constrained nonlinear problems (Cheng, Venkayya, and Khachaturian, 1976)*

The numerical method for this approach is to compute design change by using Taylor's expansions for the cost and constraint functions. Let $\mathbf{x}^{(k)}$ be the design estimate at the kth iteration and $\Delta\mathbf{x}^{(k)}$ be the change in design. Taylor's expansion of the cost and constraint functions about the point $\mathbf{x}^{(k)}$ may be expressed as
Minimize

$$F(\mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}) \cong F(\mathbf{x}^{(k)}) + \nabla F^T(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} \tag{23}$$

Subject to

$$g_j(\mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}) \cong g_j(\mathbf{x}^{(k)}) + \nabla g_j^T(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} \leq 0 \quad j = 1, \ldots, m \tag{24}$$

where ∇F and ∇g_j are gradients of the cost function and jth inequality constraint, respectively. For equality constraints, the gradients can be similarly expressed as Eq. (24). Let Eq (24) be expressed in detail as

$$\left.\begin{aligned} g_1(x) &= g_1(x^0) + \frac{\partial g_1(x^0)}{\partial x_1}(x_1 - x_1^0) + \cdots + \frac{\partial g_1(x^0)}{\partial x_n}(x - x_n) \\ &\vdots \\ g_m(x) &= g_m(x^0) + \frac{\partial g_m(x^0)}{\partial x_1}(x_1 - x_1^0) + \cdots + \frac{\partial g_m(x^0)}{\partial x_n}(x - x_n) \end{aligned}\right\} \tag{25}$$

which is expressed in matrix form as

$$[\mathrm{A}]\{x - x^0\} \leq -\{g^0\} \tag{26}$$

where $\{x^0\}$ = column matrix of the expansion point values; $\{g^0\}$ = column matrix of the nonlinear constraints evaluated at the expansion point, and $[\mathrm{A}] = \left[\frac{\partial g_i^0}{\partial x_j}\right]$; $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Let Eq. (26) be rearranged as

$$[\mathrm{A}]\{x\} \leq \{B\} \tag{27}$$

where

$$\{B\} = [\mathrm{A}]\{x^0\} - \{g^0\} \tag{28}$$

Similarly Eq. (23) can be expressed in detail as

$$\left.\begin{aligned} F(x) &= F(x^0) + \frac{\partial F(x^0)}{\partial x_1}(x_1 - x_1^0) + \frac{\partial F(x^0)}{\partial x_2}(x_1 - x_1^0) + \cdots + \frac{\partial F(x^0)}{\partial x_n}(x_n - x_n^0) \\ &= F(x^0) + \frac{\partial F(x^0)}{\partial x_1}x_1 + \frac{\partial F(x^0)}{\partial x_2}x_2 + \cdots + \frac{\partial F(x^0)}{\partial x_n}x_n \\ &\quad - \frac{\partial F(x^0)}{\partial x_1}x_1^0 + \frac{\partial F(x^0)}{\partial x_2}x_2^0 + \cdots + \frac{\partial F(x^0)}{\partial x_n}x_n^0 \end{aligned}\right\} \tag{29}$$

which in matrix form is

$$F(x) = [\mathrm{C}]\{x\} \tag{30}$$

Thus Eqs. (27) and (30) are in linear optimization form and can be solved by using standard Simplex method with iteration procedures.

4.1.1 *Optimum design of repeatedly built structures*

EXAMPLE 1(A) – DESIGN OF TWO-SPAN STEEL BEAMS

This is to show the method is very useful for designing a structure of repeat production for which efficient computer programs can be developed for daily design. A simple example of a steel highway bridge for two lane traffic is illustrated in Fig. 1 for which the load, w; span length, l_1 and l_2; allowable bending stress, F_b; and yielding stress, F_y are given. It is to determine the cross-sectional dimension. For practical construction the upper limits of flange width, b_f, flange thickness, t_f as well as the upper limit of web thickness, t_w, and web depth, d_w are required. Let design variable be

$$\{x_1 \;\; x_2 \;\; x_3 \;\; x_4\} = \{t_f \;\; b_f \;\; d_w \;\; t_w\} \tag{31}$$

then the cost function as weight of the beam may be expressed

$$F(x) = \gamma[(2b_f t_f + d_w t_w)(l_1 + l_2)] \tag{32}$$

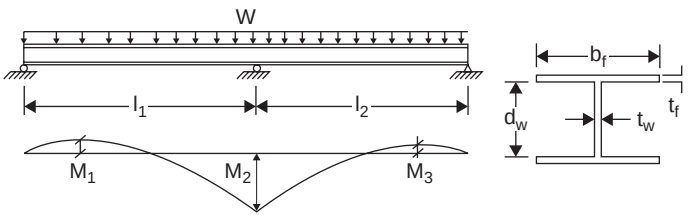


Figure 1 Example 1(A) – Two-span steel beam.

where γ is the steel weight per unit volume. From Eq. (29)

$$\left.\begin{aligned}\frac{\partial F(x^0)}{\partial x_1} &= \frac{\partial F(x^0)}{\partial t_f} = \gamma 2 b_f^0 (l_1 + l_2)\\ \frac{\partial F(x^0)}{\partial x_2} &= \frac{\partial F(x^0)}{\partial b_f} = \gamma 2 t_f^0 (l_1 + l_2)\end{aligned}\right\} \tag{33}$$

The constraints for allowable bending are expressed as

$$\left.\begin{aligned} g_i(x) &= \tfrac{M_1 c}{I_x} - F_b \le 0\\ \text{and}&\\ g_2(x) &= \tfrac{M_2 c}{I_x} - F_b \le 0\end{aligned}\right\} \tag{34}$$

Thus a typical equation in Eq. (25) is

$$g_j(x) = g_j(x^0) + \frac{\partial g_j(x^0)}{\partial x_1}(x_1 - x_1^0) + \cdots + \frac{\partial g_j(x^0)}{\partial x_4}(x_4 - x_4^0) \tag{35}$$

which involves cross-sectional properties b_f, t_f, d_w, t_w, of the derivatives of $\partial g_j / \partial x_i$. Combining all the components in Eqs. (33) and (34) yields linear optimization formulation as Eq. (27).

EXAMPLE 1(B) – DESIGN OF GABLE FRAME AND PRECAST BEAM

The idea can be extended to another repeated structural production of gable frames and precast beams as shown in Fig. 2. Note that the cost function based on the weight of steel structures is a reasonable measurement but objective function of a precast member should involve cost of concrete, steel, and construction. The objective function is usually defined as the minimization of the cost per unit surface area of the member as

$$F = \frac{1}{A}[B_1 A_c + B_2 A_s + B_3 A A_m + B_4 P_s] \tag{36}$$

where A is the flange-width of the beam; A_c is the area of the cross section; P_s is the perimeter of the section; B_1 is the cost per unit volume of concrete in place; B_2 is the

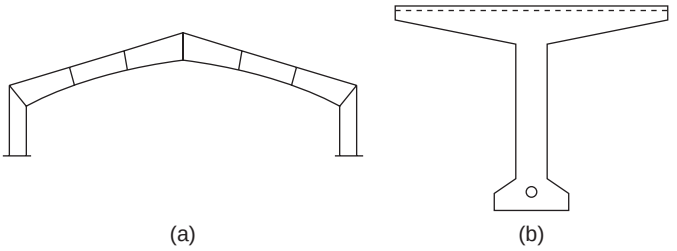


Figure 2 Example 1(B) – (a) Gable frame, (b) Precast beam.

unit cost of prestressing steel (A_s) in place; B_3 is the unit cost of mild steel A_m in place; and B_4 is a constant intended to reflect forming and finishing costs.

4.1.2 *Optimum design of frame structures*

For the algorithm, the optimization process takes a finite number of revised design parameters to achieve optimal results of a structure. Let Δz_i be the change in total weight ΔW, if W is the new total weight of the revised system, the objective function can be expressed as

$$W = W^0 + \Delta W = \sum_{i=1}^{n} \gamma L_i (z_i^0 + \Delta z_i) \tag{37}$$

where Δz_i is a variable and may be positive or negative depending on the initial assumptions; γ is the weight per unit volume; L_i is the length of member i; and n of all members of the structure. In order to ensure that the solution vector of the linear programming problem will be non-negative, we may rewrite Eq. (37) as

$$W = \sum_{i=1}^{n} \gamma L_i z_i^0 \left(1 + \frac{\Delta z_i}{z_i^0}\right) \tag{38}$$

Let $u_i = 1 + \Delta z_i / z_i^0$ in which Δz_i will be chosen to force $u_i \geq 0$. Thus, Eq. (38) becomes

$$\begin{aligned} W &= \sum_{i=1}^{n} \gamma L_i z_i^0 u_i \\ &= \gamma \sum_{i=1}^{n} L_i A_i^0 u_i = \gamma \sum_{i=1}^{n} L_i \frac{S_i^0}{K_i d_i} u_i = [C]_{n\times 1}^T \{u\}_{n\times 1} \end{aligned} \tag{39}$$

where $K = S/Ad$ which is to relate the design variable S (sectional modules) for bending and A (cross-sectional area) for axial forces; d is the cross-sectional depth. The value of K ranges from 0.32 to 0.37 for a large spectrum of rolled wide flange sections in AISC steel manual. Thus K may be assumed as 0.34 to 0.35 for practical design purposes. For the constituent members of a frame subjected to combined bending moment (M) and axial force (F), the stress constraints may be written as

$$\frac{F_{new}}{(F_{new})_{all}} + \frac{M_{new}}{(M_{new})_{all}} \leq 1 \tag{40}$$

in which F_{new} = axial force on the new section; M_{new} = bending moment on the new section; $(F_{new})_{all}$ = allowable axial force on the new section = $F_a(A + \Delta A)$; $(M_{new})_{all}$ = allowable bending moment on the new section = $F_b(S + \Delta S)$; F_a = allowable axial stress in tension or compression; and F_b = allowable bending stress. The stress constraints may be expressed in terms of cross sectional area as

$$\sum_{i=1}^{m} \left(\Delta F_{jik} + \frac{\Delta M_{jik}}{Kd F_b / F_a}\right) u_j - F_a A_j^0 u_j$$

$$\begin{cases} \leq -F_{jk}^0 - \dfrac{M_{jk}^0}{Kd F_b/F_a}, & F_{jk}^0 \geq 0 \\ \geq -F_{jk}^0 - \dfrac{M_{jk}^0}{Kd F_b/F_a}, & F_{jk}^0 < 0 \end{cases} \tag{41}$$

where j, i, and k signify the force on a member (j) due to individual member changes (i members) at the loading condition k. Eq. (41) can also be expressed in terms of section modulus as

$$\sum_{i=1}^{m} \left(\Delta M_{jik} + \frac{\Delta F_{jik}}{F_a/F_b Kd} \right) u_j - F_b S_j^0 u_j \begin{cases} \leq -M_{jk}^0 - \dfrac{F_{jk}^0}{F_a/F_b Kd}, & M_{jk}^0 \geq 0 \\ \geq -M_{jk}^0 - \dfrac{F_{jk}^0}{F_a/F_b Kd}, & M_{jk}^0 < 0 \end{cases} \tag{42}$$

The displacement constraints are formulated as

$$\left[\sum_{i=1}^{m} \Delta X_{jik} u_j \begin{array}{l} \leq (X_{pk})_{all} - 2X_{pk}^0, X_{pk}^0 \geq 0 \\ \geq (X_{pk})_{all} - 2X_{pk}^0, X_{pk}^0 < 0 \end{array} \right]; \quad \mathrm{p} = 1, np; k = 1, nc \tag{43}$$

where p signifies degree of freedom; np is the total degree of freedom including joint rotations and nodal displacements; and nc = total loading conditions. The maximum allowable deformation is usually given as nodal displacement. Because of linearization of the nonlinear problems, the side constraints are essential for move limits of solution vectors which are defined as

$$\left.\begin{array}{l} u_i \geq (LB)_i \\ u_i \leq (UB)_i \end{array}\right\} \tag{44}$$

where $(LB)_i$ and $(UB)_i$ are lower and upper bounds for solution vectors, respectively. A practical number of 0.8 (lower bound) and 1.2 (upper bound) is recommended. The primal constraints in standard form may be summarized as

$$\begin{bmatrix} \Delta F \\ \Delta X \\ I \\ I \end{bmatrix}_{nr \times n} \{u\}_{n \times 1} + [I']_{nr \times nr} \{u'\}_{nr \times 1} = \begin{Bmatrix} F' \\ X' \\ UB \\ LB \end{Bmatrix}_{nr \times 1} \tag{45}$$

nr = number of rows = $(nf)(nc) + (np)(nc) + n + n$; ΔF = matrix of coefficients in the stress constraints = $(nf \times nc) \times n$; ΔX = matrix of coefficients in the deformation

constraints $= (np \times nc) \times n$; $I =$ identity matrix $= n \times n$; $I' =$ diagonal matrix consisting of positive or negative unit coefficients corresponding to slack and surplus variables $= nr \times nr$; $u =$ solution vector of multipliers in design parameters $= n \times 1$; $u' =$ solution vector of slack and surplus variables $= nr \times 1$; $F' =$ right hand side of stress constraints $= -F^0_{jk}$, $j = 1$, nf; $k = 1$, nc; $X' =$ right hand side of deformation constraints $= (X_{jk})_{all} - 2X^0_{jk}$, $j = 1$, np; $k = 1, nc$; UB = upper bounds on the allowable percent changes in design variables $= n$ x n; LB = lower bounds on the allowable percent changes in design variables $= n$ x n; and $nf =$ number of axial forces.

The dual problem can then be written:

$$\text{Maximize } Z = \begin{bmatrix} F' \\ X' \\ UB \\ LB \end{bmatrix}^T_{1 \times nr} \{w\}_{nr \times 1} \tag{46}$$

$$\text{Subject to } \begin{bmatrix} \Delta F \\ \Delta X \\ I \\ I \end{bmatrix}^T_{n \times nr} \{w\}_{nr \times 1} \leq (C)_{n \times 1} \tag{47}$$

in which $\{w\}_{nr \times 1} =$ the dual variables; $(C)_{n \times 1} =$ coefficients in Eq. (39). At the optimum solution, the maximum value of Z should equal the minimum value of W and the solutions to the primal problem are found in the dual tableau.

EXAMPLE 2 FRAME DESIGN FOR STATIC LOADS AND WIND

Consider the single span quadrangular frame shown in Fig. 3(a) where L = 15 ft; C = 7 ft. 6 in., and D = 15 ft. Assume that girders rest at quarter points of the roof beam and may transmit vertical live load and/or dead load to the roof beam at these points. The wind load may be applied horizontally in either direction by an equivalent concentrated load at the beam-column connection. Take the values of P_1, P_2 and P_3 shown in the figure where $P_1 = P_2 = 10$ kip (dead load – D); $P_1 = P_2 = 30$ kip (live

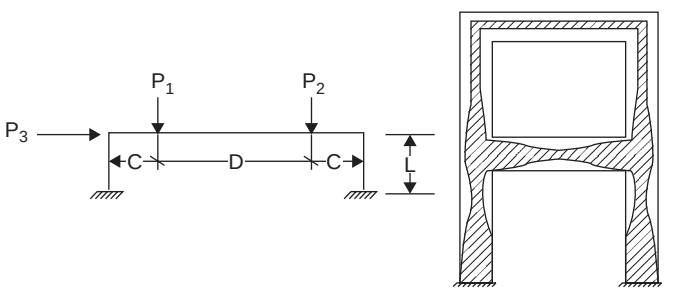


Figure 3 (a) Example 2 – Design for static load and wind; (b) Example 3 – Design for frequencies.

load – L); $P_3 = 8$ kips (wind load – W). Apply the loads in combinations shown in Eq. (48) as four loading conditions.

nc	¾(D + L + W)	¾(D + L_1 + W)	¾(D + L_2 + W)	D + L
P_1	30.	30.	7.5	40.
P_2	30.	7.5	30.	40.
P_3	6.	6.	6.	0.

(48)

The allowable stresses are assumed as $F_b = \pm 20.$ ksi, $F_a = +20.$ or $-15.$ ksi. The frame is divided into five elements for practical construction, columns (S_1 or A_1) are assumed to be identical and the girder is prismatic (S_2 or A_2); thus among five design variables, only two are independent. The solutions are given in Eq. (49) which require only three cycles for convergence (Romstad and Wang, 1968).

S	3rd cycle(in^3)
1	125.54
2	116.61

A	3rd cycle
1	20.51
2	19.05
Weight	14244.24 γ(in^3)

(49)

EXAMPLE 3 FRAME DESIGN FOR NATURAL FREQUENCIES

Design the two-story frame shown in Fig. 3(b) for the fundamental frequency $\omega \geq 24$ rad/sec. Each girder and column is divided into eight and ten elements, respectively. Half of the element mass associated with transverse motion is assumed to be lumped at the elements ends; the mass associated with element's longitudinal motion is the element's total mass. The eigenvalue solution of the frame is given as

$$[K]\{X\} = \omega^2[M]\{X\} \tag{50}$$

where [K] and [M] are system stiffness and mass matrix, respectively; {X} is eigenvector. Using linearization technique on the above nonlinear equation yields

$$[[K] + [\delta K_E]_i][\{X\} + \{\delta X\}_i] = (\omega + \delta\omega_i)^2[[M] + [\delta M_E]_i][\{X\} + \{\delta X\}_i] \tag{51}$$

in which $[\delta K_E]_i$ and $[\delta M_E]_i$ are the incremental stiffness and mass matrix of element *i*, respectively; $\delta\omega_i$ and $\{\delta X\}_i$ signify the changes of frequency and eigenvector, respectively, due to *ith* elements increment. From Eq. (51) the incremental frequency expression is given as

$$\delta\omega_i = \frac{1}{2\omega}\{X\}^T\left[[\delta K_E]_i - \omega^2[\delta M_E]_i\right]\{X\} \tag{52}$$

The cross sections from the final design are sketched in Fig. 3(b) which shows that the mass variation from the fixed bases up the columns is slightly nonlinear until the minimum area requirement is reached just below the first story level. The design of

Table 1 Data for Example 4.

Crops	*Acres per pound of seed*	*Total hours of labor per pound of seed*	*Dollar return per pound of seed*
Potatoes	0.1	0.6	4
Corn	0.2	0.5	6
Wheat	0.3	0.4	6
Hay	0.4	0.3	5

the second story is linearly varying tapers, largest in the vicinity of interconnecting members (Romstad and Runge, 1972).

EXAMPLE 4 OPTIMUM FINANCIAL MANAGEMENT

This example is purely linear optimization which is in fact originally developed for engineering management and financial investments. Let us consider a case that a farmer has a 150 acre farm. From last year's crops, he has an abundant supply of potatoes, corn, wheat, and hayseed. He wants to apportion his 150 acres among the four crops to yield a maximum return on the market. He has outside job prevents him from working more than 50 hours a week on his farm during the 10 weeks required for planting, cultivating, and harvesting. Data on the amounts of seed required, time needed to tend the crop and expected dollar return is tabulated for each crop as shown in Table 1.

With this information we can write the optimization problem in the standard linear optimization form. Let x_1, x_2, x_3, x_4 be the number of pounds of seed planted of potatoes, corn, wheat, and hay, respectively; then our linear program is

$$\left.\begin{aligned} &\text{Maximize} && F = 4x_1 + 6x_2 + 6x_3 + 5x_4 \\ &\text{Subject to} && 0.1x_1 + 0.2x_2 + 0.3x_3 + 0.4x_4 \le 150 \\ &&& 0.6x_1 + 0.5x_2 + 0.4x_3 + 0.3x_4 \le 500 \\ &&& x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0 \end{aligned}\right\} \tag{53}$$

From which the dual formation is

$$\left.\begin{aligned} &\text{Minimize} && Z = 150w_1 + 500w_2 \\ &\text{Subject to} && 0.1w_1 + 0.6w_2 \ge 4 \\ &&& 0.2w_1 + 0.5w_2 \ge 6 \\ &&& 0.3w_1 + 0.4w_2 \ge 6 \\ &&& 0.4w_1 + 0.3w_2 \ge 5 \\ &&& w_1 \ge 0, w_2 \ge 0 \end{aligned}\right\} \tag{54}$$

Note that at the point x^* that maximizes F and w^* that minimizes Z, the two functions are actually equal in value. As shown in Eqs. (39), (45), (46), and (47); the dual formation has much less number of variables. The final solution of the investment's

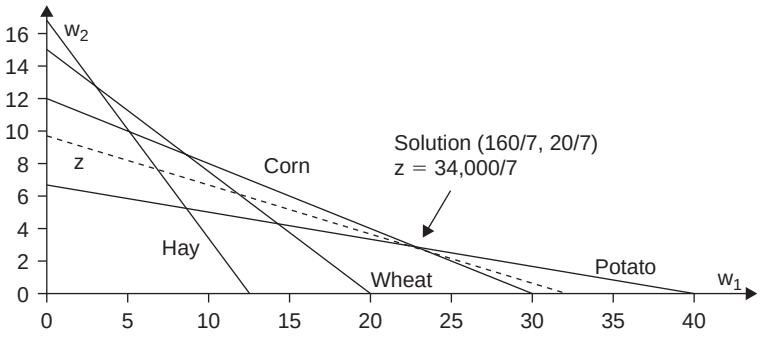


Figure 4 Example 4 – Investment Management.

problem is $w_1 = \$160/7$ per acre rent and $w_2 = \$20/7$ per hour wages which results in the total outlay of $Z = 150w_1 + 500w_2 = \$34,000/7$, as shown in Fig. 4. Note that, the Simplex method passes from vertex to vertex on the boundary of the feasible polyhedrom, repeatedly increasing the objective function until either an optimal solution is found, or it is found that no solution exists. In general engineering practice of hundreds of variables, the method is highly efficient. Problems of thousands or even more of variables can be routinely solved using the Simplex method on modern computers. The method is therefore introduced here (Cheng, Venkayya, and Khachaturian, 1976).

4.2 *Feasible direction method for constraint problem*

There are a number of mathematical programming methods for solving constraint nonlinear optimization including method of feasible directions, gradient projection method, and generalized reduced gradient method. Here a well-known method of feasible directions is introduced. The basic idea of the method is to move from one feasible design to an improved feasible design as shown in Fig. 5(a). For a given feasible design $\mathbf{x}^{(k)}$, an improving feasible direction $\mathbf{s}^{(k)}$ is determined such that using a sufficiently small step size $\alpha > 0$, the following two properties are satisfied: (i) the new design, $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \mathbf{s}^{(k)}$ is feasible, and (ii) the new cost function is smaller than the old one, i.e. $F(\mathbf{x}^{(k+1)}) < F(\mathbf{x}^{(k)})$. Once $\mathbf{s}^{(k)}$ is determined, a line search is performed to determine how far to proceed along $\mathbf{s}^{(k)}$. This leads to a new feasible design $\mathbf{x}^{(k+1)}$, and the process is repeated from there as sketched in the figure. The algorithm is to find the gradient of active constraints $\nabla g's$ and that of objective function ∇F; when they become collinear, the optimal solution is then obtained (at least the local optimum) as shown in Fig. 5(b).

EXAMPLE 5 DESIGN OF RIGID FRAMES FOR STATIC AND DYNAMIC LOAD WITH CONSIDERATION OF P-Δ EFFECTS

The 4-story rigid frame is subjected to a dynamic force, $F(t) = 10 \sin \omega t$ kips (see Fig. 6), acting on the sideways direction of the frame as shown in Fig. 7(a). The structure is also subjected to

Dead Load: 2 kip/ft; first three floors
1 kip/ft; roof

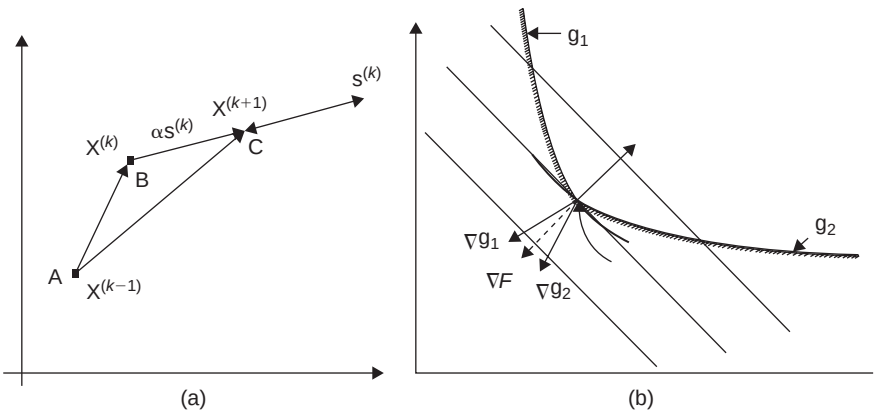


Figure 5 (a) Direction search, (b) optimal solution.

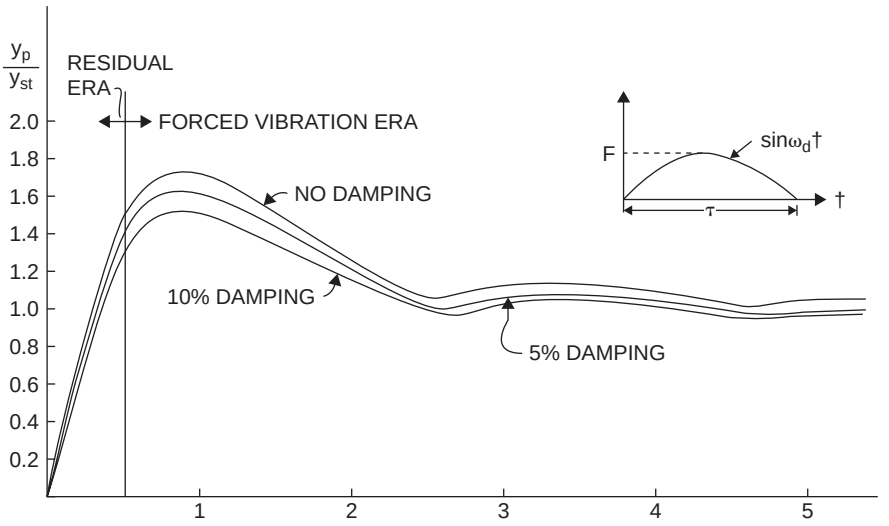


Figure 6 Dynamic force and its response spectrum.

Live Load: 1.2 kip/ft; first three floors
0.6 kip/ft; roof.

The limits on the behavior and side constraints are as follows

$Y_{pi} \leq$ [.10 .15 .20 .25] ft;	4 sideway displacements
$\sigma_{pi} \leq 20$ ksi;	24 bending stresses
5 rps $\leq p_1 \leq$ 100 rps;	fundamental frequency
10 in^3 $\leq s_l \leq$ 1000 in^3	lower and upper limits of section modulus

The rigid frame has 12 degrees of freedom and 8 design variables. The four translational degrees (sideways) of freedom are considered for eigenvectors in the dynamic analysis.

Four design cases include: (1) design based on dynamic load only with frequency constraint of $p_1 \geq 5$ rps; (2) same as (1) except the limit on the fundamental natural frequency p_1 is raised to 8 rps; (3) design based on dynamic load considering girder shears resulting from dynamic analysis which cause column axial force and then induce P-Δ effect; and (4) combined dynamic and static load design as well as P-Δ effect caused by both loads.

The P-Δ effect is considered for columns of which each member is subjected to axial force, P, resulting from static and dynamic analysis. For the dynamic force of half sine wave, a shock spectrum is also given in Fig. 6 where y_p/y_{st} represents the ratio of dynamic displacement, y_p, to static displacement, y_{st}. Using modal matrix in the dynamic analysis, one may find the peak response as

$$Y_{pi} = \sum_{r=1}^{nps} \left| X_{ir} c_r y_{pr} \right| \tag{55}$$

in which X_{ir} is the ith component of the rth eigenvector; nps is total number of modes corresponding to the number of sideways; $c_r = \{X\}^{(r)T}\{F_0\}$; $F_0 = 10$ kips; y_{pr} is the peak response from the spectrum at rth frequency. Similarly the peak stress at a given member's end can be written in component form as

$$\sigma_{pi} = \sum_{r=1}^{nps} \left| \sigma_{bir} c_r y_{pr} \right| \quad i = 1, \ldots, 2n \tag{56}$$

where σ_{bir} is the ith component of $\{\sigma_b\}^{(r)}$ and n is the number of members.

A plot of the cumulative number of design cycles versus the weight of the design for all four cases is shown in Fig. 7(b). It can be seen that by neglecting the effects of the girder shears from dynamic analysis, the optimal weight is 6.050 kips as shown in curve (1). Approximately a 10% heavier design is produced if girder shears from the dynamic analysis are included in the design as shown in curve (3). Curve (4) indicates that if P-Δ effects due to static and dynamic forces are considered, then the final design weight (8.007 kip) is increased markedly. Curve (2) shows that imposing a higher restriction on the fundamental frequency results in a heavier final design. The restriction can be important if the structure supports a machine operating in the neighborhood of the natural frequency (Cheng and Botkin, 1976; Cheng, Jiang, and Lou, 2008).

4.3 *Transformation method*

The transformation method is to solve constrained optimization problem by using non-constrained optimization technique. The basic idea is to construct a composite function using objective function and constraint function with certain parameters (called penalty parameters). Therefore the transformation method can also be called penalty function method such as interior penalty function method and exterior penalty function method, among others.

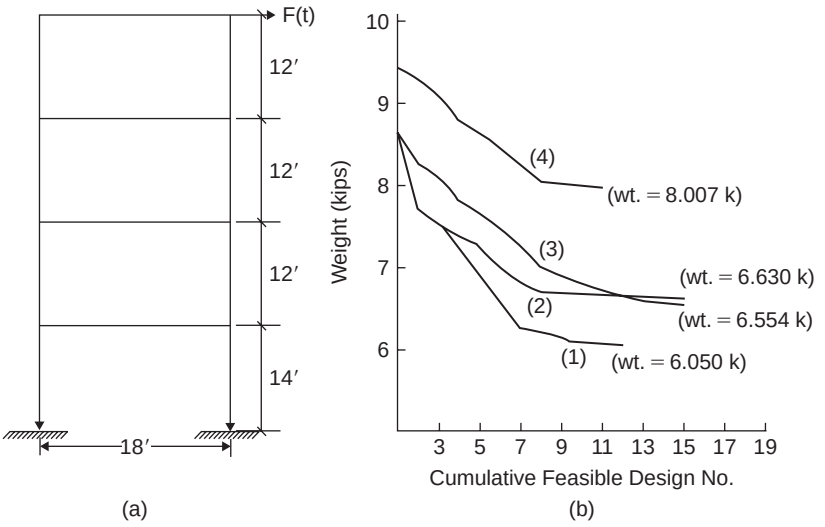


Figure 7 Example 5 – (a) Given frame; (b) Optimal weight.

The penalty function is a procedure for approximating constrained optimization problem. The approximation is accomplished by adding to the objective function – a term that prescribes a high penalty value for violation of constraints. The penalty parameter, r, determines the severity of the penalty and, consequently, the degree to which the unconstrained formulation approximates the original constrained problem. The typical equations of exterior and interior penalty functions are given in Eqs. (57) and (58), respectively.

$$\psi(x,r) = F(x) + r\sum_{j=1}^{m} (g_j(x))^2 \tag{57}$$

$$\psi(x,r) = F(x) - r\sum_{j=1}^{m} \frac{1}{g_j(x)} \tag{58}$$

Let us examine Eq. (58) where F is to be minimized over all x satisfying $g_j(x) < 0$, $j = 1, 2, \ldots, m$. Note that if r is positive the effect is to add a positive penalty to $F(x)$. This is because at an interior point, all the terms in the sum are negative. As a boundary is approached, some g_j will approach zero and the penalty will blow up. The penalty parameter r will be made successively smaller in order to obtain the constrained minimum of F. similar arguments can be made for exterior penalty function given in Eq. (57). A simple diagram of convergent behavior of these two methods are sketched in Fig. 8 based on two constraints, $g_1(x)$ and $g_2(x)$, as well as two design variables, x_1 and x_2; where $\boldsymbol{x}^*$ is the optimal solution.

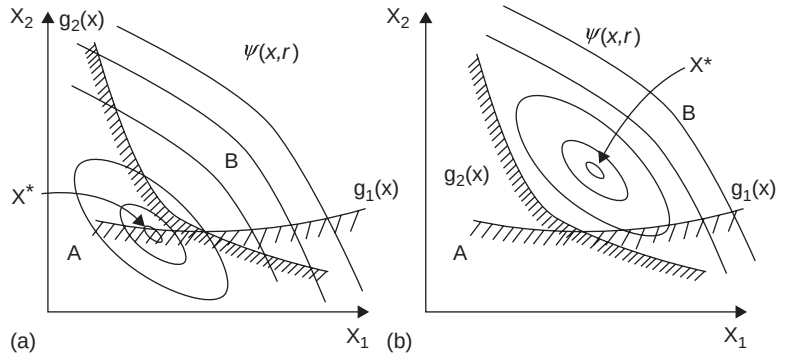


Figure 8 Illustration of penalty function method (a) Exterior Penalty; (b) Interior penalty.

EXAMPLE 6 INTERIOR PENALTY FUNCTION METHOD FOR AN ALGEBRA PROBLEM

$$\text{Minimize} \quad x_1 + x_2 \tag{59}$$

$$\text{Subject to} \quad \left.\begin{aligned} g_1 &= 3 - x_1 < 0 \\ g_2 &= 2 - x_2 < 0 \end{aligned}\right\} \tag{60}$$

For interior penalty function method, the penalty function is

$$\psi(x, r) = x_1 + x_2 + r\left(\frac{1}{3 - x_1} + \frac{1}{2 - x_2}\right) \tag{61}$$

Using $\partial\psi/\partial x_1 = 0, \partial\psi/\partial x_2 = 0$ yields the minimum solution of ψ. Applying these two conditions to Eq. (61) gives

$$\left.\begin{aligned} x_1 &= 3 - r^{1/2} \quad \text{or} \quad 3 + r^{1/2} \\ x_2 &= 2 - r^{1/2} \quad \text{or} \quad 2 + r^{1/2} \end{aligned}\right\} \tag{62}$$

From which the values $x_1 = 3 + r^{1/2}$ and $2 + r^{1/2}$ are the feasible points for the desired solution. As r approaches to zero, the optimum result is 5 at $x_1 = 3$ and $x_2 = 2$. Observing Eq. (61) indicates that when the constraint approaches to the boundary, then the penalty blows up and r should be decreased to zero.

As shown in Fig. 8, the exterior penalty function method is to search the solution in the infeasible region (A) of which the advantage is that solution may be started from any infeasible point. The disadvantage is due to the fact that only the optimal solution can be used. The interior penalty function method, however, is to search the solution in the feasible region (B), therefore the advantage is that the solutions are always feasible. But the starting point should be carefully selected in order to be in the feasible region. It is worthwhile noting that some optimization problems are difficult, if not impossible, to express derivatives of constraints with respect of design variables

as described previously. This type of problem is relatively easier to be solved by using the penalty function method.

A typical problem is the optimal design of nondeterministic structures subjected to dead load, live load and dynamic forces with various considerations of probability concepts of variance formulations as well as loading models. Two numerical examples are to show the complexity of this type of optimization problem. Before numerical illustrations, some essential formulations are briefly reviewed as follows (Ang and Tang, 2007):

A) Safety factor expressions

The probabilities of failure may be expressed as

$$P_f = P(R < S) = \int_0^\infty F_R(r) f_s(r)\, dr \tag{63}$$

in which R = resistance, S = response, r = random parameters, F_R = cumulative probability distribution function for a resistance, and f_s = probability density function for a response. If the probabilities of failure are assumed to have normal distribution, they may be expressed as

$$P_f = 1 - \Phi(\beta) \tag{64}$$

in which $\Phi()$ = standard normal probability distribution and β = safety factors. Depending on the relationship between responses and resistances, safety factors may have normal or lognormal distribution with the following expressions (Ellingwood and Ang, 1972):
normal (N),

$$\beta = \frac{\overline{R} - \overline{S}}{(\sigma_R^2 + \sigma_S^2)^{\frac{1}{2}}} \tag{65}$$

lognormal (LN),

$$\beta = \ell_n \left[\frac{\overline{R}}{\overline{S}\sqrt{\frac{1+V_S^2}{1+V_R^2}}} \Big/ \sqrt{\ell_n[(1+V_R^2)(1+V_R^2)]} \right] \tag{66}$$

in which $\overline{R}$, $\overline{S}$ = mean of resistance and response, respectively; σ_R^2, σ_S^2 = variance of resistance and response, respectively; and V_R, V_S = coefficient of variation of resistance and response, respectively. Because the functional relationship between the response (or resistance) and the parameters is difficult to determine, first order approximation can be used to find the mean and variance of response (resistance) about the mean parameter values, i.e.,

$$\overline{S}(r) = S(\overline{r}), \quad \overline{R}(r') = R(\overline{r}') \tag{67}$$

$$\sigma_S^2(r) = \sum_i \sum_j \left(\frac{\partial S}{\partial r_i}\right)_{\bar{r}} \left(\frac{\partial S}{\partial r_j}\right)_{\bar{r}} \rho_{r_i r_j} V_{r_i} V_{r_j} \bar{r}_i \bar{r}_j \tag{68}$$

$$\sigma_R^2(r') = \sum_i \sum_j \left(\frac{\partial S}{\partial r'_i}\right)_{\bar{r}'} \left(\frac{\partial S}{\partial r'_j}\right)_{\bar{r}'} \rho_{r'_i r'_j} V_{r'_i} V_{r'_j} \bar{r}'_i \bar{r}'_j \tag{69}$$

in which $r(r')$ = random parameters of response (resistances), $\rho_{r_i r_j}(\rho_{r'_i r'_j})$, $V_{r_i}(V_{r_i})$, $V_{r_j}(V_{r_j})$, $\bar{r}_i(\bar{r}_i)$, $\bar{r}_j(\bar{r}_j)$ = the correlation coefficients, coefficients of variation, and mean values of ith or jth random parameters of the responses (resistances), respectively.

B) Loads and Load Effects

The applied loadings in structures can be dead (D) or live (L) or earthquake (E) loads or a combination of them. The load effects mean the responses (i.e. displacements or internal forces) caused by these loadings. The determination of these loads and their effects should include uncertainties. In the first numerical example, four models of mean ($\overline{L}$) and their associated coefficients of variation (V_L^2) of live load are used as the US models L_1, L_2 (NBS 577, 1980; Ellingwood and Culver, 1977), English model L_3 (Mitchell and Woodgate, 1971), and deterministic model L_4. For the second example, the earthquake forces are based on the Newmark's spectra.

C) Optimization Formulation

The optimization problem is formulated as:

$$\text{Minimize} \quad W(x_1, x_2, \ldots, x_i); \quad i = 1, \ldots, n \tag{70}$$

$$\text{Subject to} \quad g_j(x_1, x_2, \ldots, x_i) \leq 0; \quad j = 1, \ldots, m \tag{71}$$

For considering structural safety, the constraints may be expressed in terms of safety factor, β, or failure probability, $P_f(\beta)$. Illustrations of constraints in terms of safety are given as follows.

For displacements of a structural system:

$$\beta_{uojd} - \beta_{ujd} \left(\frac{(\overline{u}_{ojd} - \overline{u}_{jd})}{(\sigma_{uojd}^2 + \sigma_{ujd}^2)^{\frac{1}{2}}}\right) \leq 0 \tag{72}$$

For moment resistance of beams:

$$\beta_{bojb} - \beta_{bjb} \left(\frac{(\overline{M}_y - \overline{M})}{(\sigma_{M_y}^2 + \sigma_M^2)^{\frac{1}{2}}}\right) \leq 0 \tag{73}$$

For the resistances of columns including a) yielding, b) instability in the plane of bending, c) lateral torsional buckling, and d) buckling about weak axis:

$$\beta_{f1ojc} - \beta_{f1jc} \left(\frac{(1.0 - \overline{f}_{1jc})}{(\sigma_{f1jc}^2)^{\frac{1}{2}}}\right) \leq 0 \ldots\ldots (a) \tag{74}$$

$$\beta_{f2ojc} - \beta_{f2jc}\left(\frac{(1.0-\overline{f}_{2jc})}{(\sigma^2_{f2jc})^{\frac{1}{2}}}\right) \leq 0 \ldots\ldots (b) \tag{75}$$

$$\beta_{f3ojc} - \beta_{f3jc}\left(\frac{(1.0-\overline{f}_{3jc})}{(\sigma^2_{f3jc})^{\frac{1}{2}}}\right) \leq 0 \ldots\ldots (c) \tag{76}$$

$$\beta_{f4ojc} - \beta_{f4jc}\left(\frac{(1.0-\overline{f}_{4jc})}{(\sigma^2_{f4jc})^{\frac{1}{2}}}\right) \leq 0 \ldots\ldots (d) \tag{77}$$

in which jd = jdth displacement constraint, jb = jbth beam constraint, jc = jcth column constraint, β_{uojd}, β_{bojb}, β_{f1ojc}, β_{f2ojc}, β_{f3ojc}, β_{f4ojc} = specified allowable safety factors, u_{jd} = displacement, u_{ojd} = allowable displacement, M = applied moment at the end of a member, M_y = yielding moment of a member, $f_1 = P/P_y + M/M_y$, $f_2 = P/P_{cr} + C_m M/[M_y(1- P/P_E)]$, $f_3 = P/P_{cr} + C_m M/[M_{cr}(1 - P/P_E)]$, $f_4 = P/P_{cry}$, P = applied axial load, P_y = yielding axial load, P_{cr} = critical axial load, P_E = Euler buckling load, C_m = a coefficient to account for the boundary condition of a member. The determination of the statistics of f_1, f_2, f_3, and f_4 may be obtained from design specifications.

EXAMPLE 7 DESIGN OF NONDETERMINISTIC STRUCTURES FOR DEAD AND LIVE LOADS (CHENG AND CHANG, 1988; CHANG, GER, AND CHENG, 1994)

The two story steel frame shown in the Fig. 9(a) is used to illustrate the optimum design results shown in Fig. 9(b) and (c) based on normal (N) and lognormal (LN), respectively. The notations in the figures are: L_1, L_2, L_3 and L_4 signify the live load models, 1st and 2nd represent the two variance expressions. All the designs are due to D + L (dead + live). The parameters used in the example are: $A_I = 900\ ft^2$, D = 80 psf, $V_D = 0.12$; allowable individual failure probabilities = 0.0001, allowable joint rotations = 0.05 rad, and displacements = 0.5 in, allowable variances of joint rotations and displacements all assumed to be zero, $F_y = 36$ ksi, $E = 29000$ ksi, $V_{M_y} = 0.12$, $V_{PE} = 0.3$, $V_{P_{cr}} = 0.31$, $V_{P_y} = 0.14$, $V_{M_{cr}} = 0.20$, $V_{P_{cry}} = 0.30$, $\rho_{P_y M_y} = 0.8$, $\rho_{P_{cr} M_y} = 0.8$, $\rho_{P_{cr} P_E} = 0.8$ (=1. if $P_{cr} = P_E$), $\rho_{P_E M_y} = 0$, $\rho_{P_{cr} M_{cr}} = 0.15$, $\rho_{P_E M_{cr}} = 0.15$, and Cm = 1.0. The results shown in the figure indicate that the magnitude of optimum weight is in the order of $L_4 > L_3 > L_1 > L_2$ for both normal and lognormal distributions. However, the lognormal distribution requires a heavier structural design than the normal. The 2nd variance expression yields a heavier structural design for the lognormal distribution. The English model (L_3) demands a heavier structural design than the U.S. models (L_1 and L_2). L_4 demands the heaviest simply because it does not have the effect of the influence area, A_I.

EXAMPLE 8 DESIGN OF NONDETERMINISTIC STRUCTURE FOR NONDETERMINISTIC RESPONSE SPECTRA

The ten story framework shown in Fig. 10(a) is used to illustrate (a) design for horizontal earthquake force (H) only, and (b) design for horizontal (H) and vertical (V)

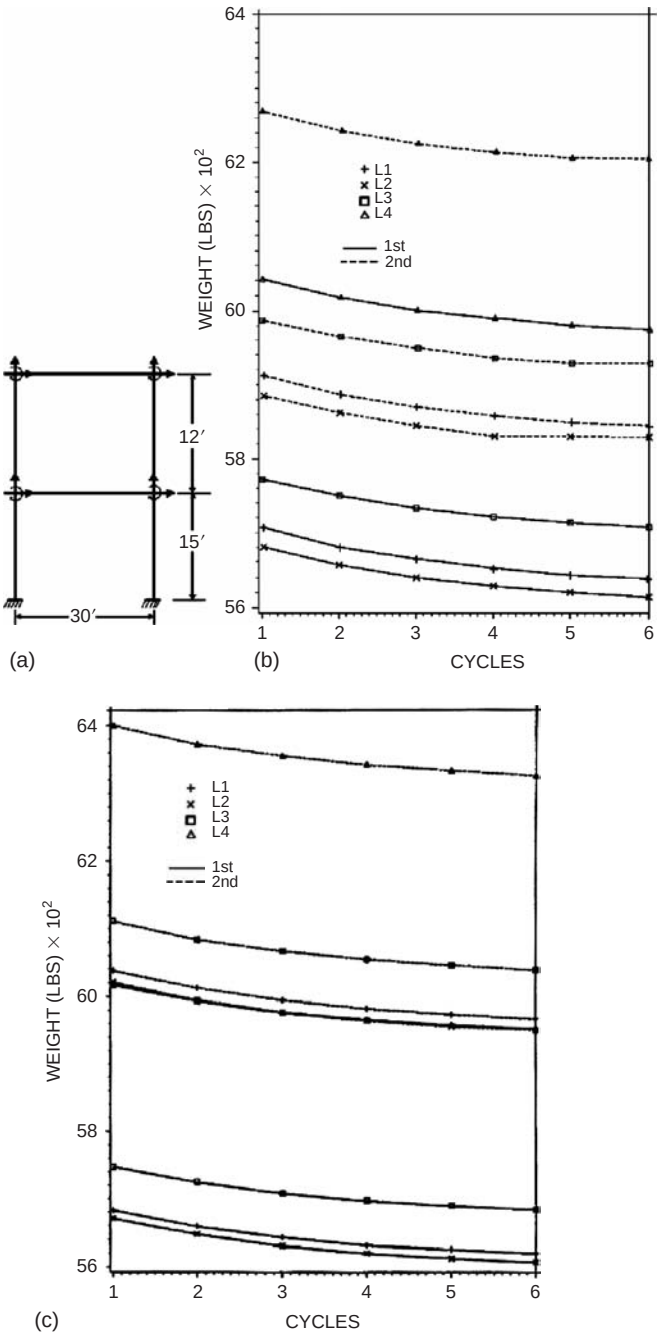


Figure 9 Example 7 – (a) Given frame, (b) D + L for N model, (c) D + L for LN model.

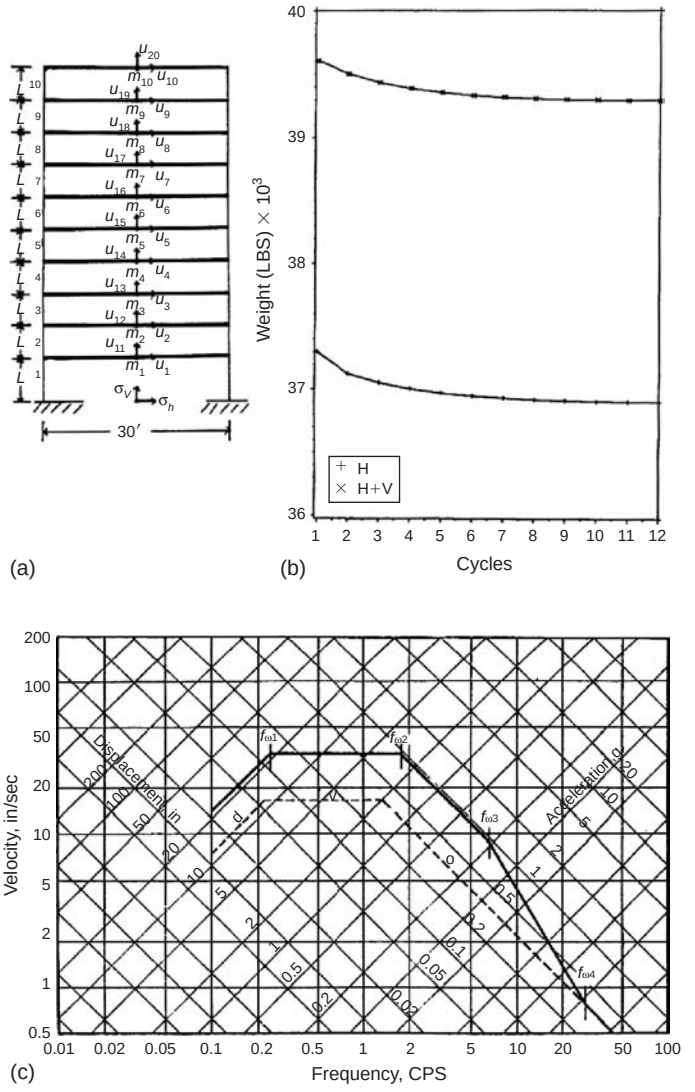


Figure 10 Example 8 – (a) 10-story frame, (b) Optimum weight, (c) Newmark's NNSRS.

ground motions. The results are shown in Fig. 10(b). The Natham Newmark's nondeterministic response spectrum (NNSRS) given in Fig. 10(c) is used for the earthquake excitation (Mohraz, Hall, and Newmark, 1972). The uniform weight for each story is 0.27 k-sec^2/in. The ratio of vertical to horizontal ground acceleration is assumed to be 4/3. For the response spectrum, the statics of displacement, the mean, the variance, and the coefficient of variation of spectral displacement for each mode are derived for different frequency ranges: $\overline{f}_{\omega_1} > \overline{\omega}_n/2\pi, \overline{f}_{\omega_1} > \overline{\omega}_n/2\pi < \overline{f}_{\omega_2}, \overline{f}_{\omega_2} > \overline{\omega}_n/2\pi < \overline{f}_{\omega_3}$, and

$\bar{f}_{\omega_4} < \bar{\omega}_n/2\pi$. The seismic displacement response is based sum of square of displacement corresponding to each mode and the 1st variance and 2nd variance are also derived (Cheng and Chang, 1988). The mean and variance of structural resistance are the same as for static load as discussed before. The failure modes as constraints are yielding failure, lateral buckling failure, and torsional buckling failure. The P-Δ effects due to vertical forces are not considered. As shown in Fig. 10(b) the optimum solution against cycle for horizontal or combination of horizontal and vertical ground accelerations based on normal (N) and second variance approach indicate that horizontal coupled with vertical ground accelerations demand heavier structural design.

5 Optimality criterion methods

It has been commonly recognized that the mathematical programming techniques, as discussed in previous sections, are general in application but quite restricted in the number of design variables. Herein the optimality criterion techniques are introduced which can handle a large number of design variables but lack generality. Consequently a considerable amount of research interests have been undertaken for achieving more engineering applications. The optimality criterion techniques have been developed in several broad categories including: 1) envelope optimality criterion technique, 2) recursive optimality criterion technique, 3) generalized optimality criterion technique, and 4) computer aided design of building optimization. These techniques are sequentially discussed in the following sub-sections. The optimality criterion can be generally stated as the optimum structure is one in which the average strain energy density when combined with the kinetic energy density is the same for all of the constituent members. The statement can be accepted for single constraint, single loading structures. For more complicated structures, the optimality criterion is not so simply stated, but it is explicitly or implicitly defined in mathematical or numerical operations (Khot, Berke, and Venkayya, 1979; Venkayya, Khot, and Reddy, 1969).

5.1 *Envelope Optimality Criterion Technique*

Let Eqs. (5) – (8) be written as

$$\frac{\partial W(x)}{\partial x_i} + \sum_{j=1}^{M} \lambda_j \frac{\partial g_j(x)}{\partial x_i} = 0, \quad i \in J \tag{78}$$

and

$$\frac{\partial W(x)}{\partial x_i} + \sum_{j=1}^{M} \lambda_j \frac{\partial g_j(x)}{\partial x_i} \geq 0, \quad i \in J_0 \tag{79}$$

in which $W(x)$ is explicitly expressed in terms of structural weight; J is the set of design variables for the conditions of $x_i > x_i^0$; J_0 is the set of design variables constrained by the lower bounds, x_i^0; and M denotes the number of active behavior constraints. Let

$$\mu_{ji} = \frac{\partial g_j(x)}{\partial x_i}, \quad \tau_i = \frac{\partial W(x)}{\partial x_i} \tag{80a,b}$$

The KT condition yields for M active behavior constraints as

$$-\frac{\sum_{j=1}^{M} \lambda_j \mu_{ji}}{\tau_i} = 1, \quad i = 1, 2, \ldots, n. \tag{81}$$

Optimality criterion method has an important numerical procedure called iteration and scaling factor which are developed for saving computational efforts. Let the scaling factor be Λ which expresses the design variable x_i in terms of relative design variable α_i as

$$x_i = \Lambda \alpha_i, \quad i = 1, 2, \ldots, n \tag{82}$$

Eq. (81) may be rederived for iteration between k and $k+1$ cycles as shown in Eq. (83) with a square root is a convergence parameter.

$$(\Lambda \alpha_i)^{(k+1)} = (\alpha_i)^{(k)} \left[-\frac{\sum_{j=1}^{M} \lambda_j \mu'_{ji}}{\tau'_i} \right]^{(k)1/2}, \quad i = 1, 2, \ldots, n \tag{83}$$

where

$$\mu'_{ji} = \Lambda \mu_{ji}, \quad \tau'_i = \frac{\tau_i}{\Lambda} \tag{84a,b}$$

5.1.1 *Calculation of lagrange multipliers for multiple active constraints*

When the problem has a single active constraint, Eq. (83) can be conveniently employed for optimum solutions. For multiple active constraints, there are a few developments pertinent to the optimality criterion formulations. The calculation of the multipliers for Eq. (83) is to form an auxiliary function, $L(\lambda)$, as follows

$$l(\lambda) = \sum_{i=1}^{n} \left[1 + \frac{\sum_{j=1}^{M} \lambda_j \mu_{ji}}{\tau_i} \right]^2 \tag{85}$$

which then is minimized by solving the set of linear equations for the Lagrange multipliers, λ_j, as

$$\frac{\partial L(\lambda)}{\partial \lambda_j} = 0 \quad j = 1, 2, \ldots, M \tag{86}$$

5.1.2 *Constraints formulations*

The behavior constraints, $g_j(x)$, considered in this approach has separate formulations in terms of limitations of displacements (based on flexibility) and stresses (based on

stiffness) that are imposed on both static and seismic structures. However, for seismic structures, an additional behavior constraint of natural frequencies is taken into account (Cheng and Srifuengfung, 1978).

A. FLEXIBILITY CONSTRAINTS FOR STATIC LOADS

The displacement constraint function may be expressed by using the following virtual work at any nodal point:

$$u_j(x) = \{Q_j\}^{\mathrm{T}}\{r(x)\} \tag{87}$$

in which $\{Q_j\}$ = load vector with unit value for the jth direction and zero values for others, and $\{r(x)\}$ = vector of generalized displacements attributable to the static load, {R}. Let $\{q_j(x)\}$ be a vector of the generalized displacements attributable to load $\{Q_j\}$, then

$$\{q_j(x)\}^{\mathrm{T}} = ([K]^{-1}\{Q_j\})^{\mathrm{T}} = \{Q_j\}^{\mathrm{T}}[K]^{-1} \tag{88}$$

where $[\mathrm{K}]^{-1}$ is inversion of the system matrix [K]. $[\mathrm{K}] = [\mathrm{K_E}] + [\mathrm{K_G}]$ composed of system's stiffness matrix $[\mathrm{K_E}]$ and geometric matrix $[\mathrm{K_G}]$ (due to P-Δ effect including structural and nonstructural mass). Derivative of Eq. (87) with respect to i yields

$$\frac{\frac{\partial g_j(x)}{\partial x_i}}{\frac{\partial W(x)}{\partial x_i}} = \frac{\{q_j(x)\}^{\mathrm{T}}[K_{Ei}]\{r(x)\} - P'_i\{q_j(x)\}^{\mathrm{T}}[K_{Gi}]\{r(x)\}}{\rho_i\eta_i x_i \ell_i}$$
$$= \text{constant}, \quad i = 1, 2, ..., n \tag{89}$$

in which $[\mathrm{K_{Ei}}]$ and $[\mathrm{K_{Gi}}]$ are stiffness and geometric matrix of element i resulting from $\partial[K_E]/\partial x_i$ and $\partial[K_G]/, \partial x_i$, respectively; η_i is the ratio of cross-sectional area A_i, to the moment of inertia, x_i (design variables for beams and columns), for ith element; $P'_i = \rho_i\eta_i\ell_i/2$ (P-Δ effect due to structural mass), and ρ_i is mass density of element i. Eq. (89) may reveal some physical meaning that the optimum structure for the specified displacement is the one in which the ratio of the average virtual strain energy density to the mass density is the same for all its members.

B. STIFFNESS CONSTRAINTS FOR STATIC LOADS

The stiffness constraints are used to measure the limitations of the allowable shear stress and the allowable combined stress of axial and bending. The stiffness of the structure can be described by the work caused by the static load, {R}, multiplied by the generalized displacement, $\{r(x)\}$, in the form of

$$z(x) = \frac{1}{2}\{R\}^{\mathrm{T}}\{r(x)\} \tag{90}$$

because the product, $\{R\}^{\mathrm{T}}\{r(x)\}$, is an inverse measure of the stiffness. Thus, $z(x)$ may be called a measuring function of static stiffness. The optimality criteria for a single loading condition can be similarly derived as shown in Eq. (89) and the result is

$$\frac{1}{2}\frac{\{r(x)\}[K_{Ei}]\{r(x)\} - P'_i\{r(x)\}[K_{Gi}]\{r(x)\}}{\rho_i\eta_i x_i \ell_i} = \text{constant} \quad i = 1, 2, \ldots, n \tag{91}$$

Note that numerator of Eq. (91) represents the average strain energy in which the second term is associated with the P-Δ effects of column members only. Thus a general statement for stress constraint can be similarly made as given in (A).

C. FLEXIBILITY CONSTRAINTS FOR DYNAMIC LOADS

The dynamic displacement constraint function can be expressed in a form similar to that of Eq. (87) in terms of virtual work as follows:

$$u_j(x,t) = \{Q_j\}^{\mathrm{T}}\{r(x)\} \tag{92}$$

in which $\{Q_j\}$ = load vector with unit force for the jth direction only, and the time function is also a unit value, and $\{r(x,t)\}$ = vector of generalized displacements attributable to the dynamic load, $\{\mathrm{R}(t)\}$. Let $q_{\mathrm{j}}(x,t)$ be a vector of the generalized displacements attributable to load $\{Q_j\}$, then

$$\{q_j(x,t)\} = [\Phi][\psi][D][\Phi]^{\mathrm{T}}\{Q_j\} \tag{93}$$

where $[\Phi]$ represents the normal mode eigenvectors; $[\psi]$ is a diagonal matrix in which $\psi_i = 1/m_i\omega_i$; and [D] is also a diagonal matrix composed of Duhemel's integrals associated with individual natural frequency, ω_i. The derivative of displacement constraints may be expressed as

$$\begin{aligned}\frac{\partial u_j(x,t)}{\partial x_i} = -\frac{1}{x_i}[&\{q_j(x,t)\}^{\mathrm{T}}[K_{Ei}]\{r(x,t)\} - P_i'\{q_j(x,t)\}^{\mathrm{T}}[K_{Gi}]\{r(x,t)\} \\ &- p^2\{q_j(x,t)\}^{\mathrm{T}}[M_{Ei}]\{r(x,t)\}]\end{aligned} \tag{94}$$

where p is the frequency of the vibrating system and is obtained by using Rayleigh quotient.

D. STIFFNESS CONSTRAINTS FOR DYNAMIC LOADS

The measuring function of dynamic stiffness may be established in a manner similar to that of Eq. (90) as follows

$$z(x,t) = \frac{1}{2}\{R_0\}^{\mathrm{T}}\{r(x,t)\} \tag{95}$$

in which $\{R_0\}$ represents the magnitude of the dynamic load vector. Following Eqs. (93) and (94), we can obtain the derivative of dynamic stiffness constraint as follows

$$\{r(x,t)\} = [\Phi][\psi][D][\Phi]^{\mathrm{T}}\{R_0\} \tag{96}$$

$$\begin{aligned}\frac{\partial z(x,t)}{\partial x_i} = \frac{1}{2x_i}[&\{r(x,t)\}^T[K_{Ei}]\{r(x,t)\} + \{r(x,t)\}[K_{Gi}]\{r(x,t)\} \\ &- p^2\{r(x,t)\}[M_{Ei}]\{r(x,t)\}]\end{aligned} \tag{97}$$

Eq. (97) represents the average strain energy density combined with kinetic energy density of any member, i.

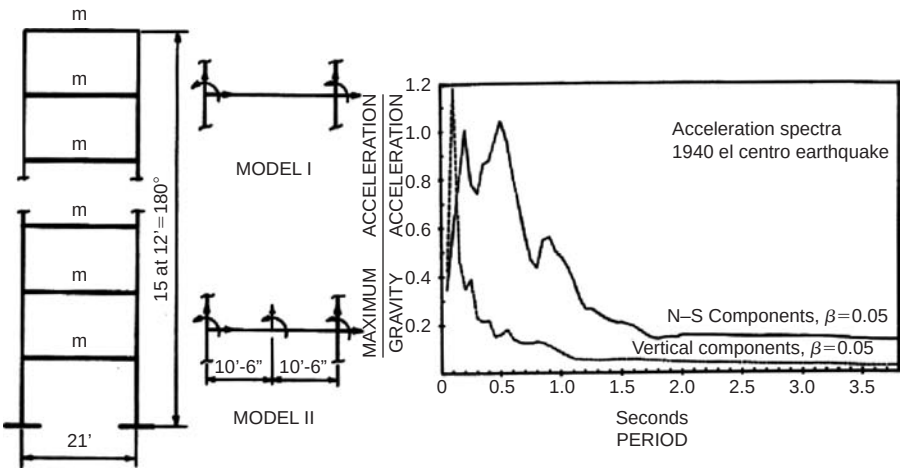


Figure 11 Example 8 – (a) 15-story frame models (b) Acceleration spectra.

E. CONSTRAINTS IMPOSED ON THE NATURAL FREQUENCIES

The natural frequency of any mode, ω_j, of a structure can be expressed as

$$\omega_j^2 = \frac{\{\Phi_j\}^T[K]\{\Phi_j\}}{\{\Phi_j\}^T[M]\{\Phi_j\}} \tag{98}$$

in which ω_j represents not only the natural frequency but also a mathematical function. Thus the derivation of the function is

$$\frac{\partial \omega_j^2}{\partial x_i} = \frac{\{\Phi_j\}^{\mathrm{T}}[K_{Ei}]\{\Phi_j\} - P_i'\{\Phi_j\}^{\mathrm{T}}[K_{Gi}]\{\Phi_j\} - \omega_j^2\{\Phi_j\}^{\mathrm{T}}[M_{Ei}]\{\Phi_j\}}{\rho_i \eta_i x_i \ell_i}$$
$$= \text{constant}, \quad i = 1, 2, \ldots, n \tag{99}$$

Note that the numerator of the above equation represents the difference between the strain energy density and the kinetic energy density of the jth mode. The second term of the numerator is resulting from compressive force on the columns.

EXAMPLE 8 DESIGN OF 15-STORY FRAME FOR MULTI-COMPONENT SEISMIC MOTIONS

The 15-story steel buildings shown in Fig. 11(a) is considered to have been subjected to the horizontal and vertical ground motions of the 1940 El Centro earthquake for which the acceleration spectra with 5% damping are given in Fig. 11(b). In addition to the consideration of vertical component of the ground motion, the second-order P-Δ effect on the design is also included. The P-Δ effect is a result of all axial forces exerted on the columns. These forces are composed of the dead loads of the structural and nonstructural masses and the associated inertial forces occasioned by vertical acceleration. Note that the structural formulation is based on the consistent mass method

Table 2 Final Weights, Natural Periods, and Displacements of 15-Story Frames (A = with Stress Constraints, B = with Stress and Displacement Constraints.

Group	*Case*	*Final Weight (kips)*	*Natural Period (sec.)*					*Disp. at top floor (in.)*
			1	*2*	*3*	*4*	*5*	
A	a	51.88	2.409	0.796	0.462	0.327	0.245	13.19
	b	53.03	2.412	0.792	0.461	0.323	0.243	13.21
	c	52.15	2.413	0.792	0.461	0.326	0.245	13.22
	d	53.32	2.409	0.786	0.458	0.322	0.242	13.22
B	a	60.44	2.152	0.694	0.404	0.282	0.212	10.80
	b	62.21	2.179	0.671	0.386	0.268	0.202	10.80
	c	60.77	2.158	0.692	0.402	0.281	0.217	10.80
	d	62.21	2.152	0.677	0.395	0.276	0.212	10.80

(Cheng, 2001). The structure is assumed to have two models: Model I has two nodes at both ends of a girder, and Model II has three nodes at both ends and at the mid-span of a girder.

The given data are: the span length and floor height are 21 ft. and 12 ft., respectively. The dead load (nonstructural mass) on each floor is 180 lbs/in. the mass density and the modulus of elasticity of the construction material are 0.283 lbs/in^3 and 29,000 ksi, respectively. The allowable stress, σ, for bending combined with the axial forces of the columns and girders is 29 ksi, and the allowable shear stress, σ_v, is 0.65σ. Although different allowable deflections may be imposed at any particular node, the allowable deflection of each floor is limited to 0.005 times the height of that floor from ground level. The columns and girders are made of the built-up sections have the following properties: $b_f = 25$ in., $(d_w)_{max} = 75$ in., $(d_w)_{min} = 15$ in., $t_f/(d_w)_{max} = 0.045$, $t_f/(d_w)_{min} = 0.023$. Four design cases are: Case (a) for horizontal ground motion only (H), Case (b) for a horizontal ground motion and the P-Δ effect of the dead load associated with the structural and nonstructural masses (H + PΔ(DL)), Case (c) for horizontal and vertical earthquake components but no P-Δ effect (H + V), and Case (d) for horizontal and vertical earthquake components as well as the P-Δ effect occasioned by the dead load and the vertical inertial forces associated with the structural and nonstructural masses (H + V + PΔ + (DL + V)) (Cheng and Srifuengfung, 1978).

Table 2 lists the final weights of these four cases. The table also includes the final displacement at the top floor, the number of modes, and the associated natural periods used in the design. It is apparent that the multi-component ground motion when combined with the P-Δ effect of the structural and nonstructural masses can yield nearly 3% of an increase in the structural weight over that which would be required for one horizontal component only.

Fig. 12 shows the ratio of the energy (kinematic and strain), W, of the individual modes to the total energy, W_T, associated with the total number of natural modes included in the design. This plot signifies that the first mode is the most significant for all cases. The modes beyond the third have little effect on a structure subject to a horizontal motion such as in Cases (a) and (b) of Model I. However, when a structure is subjected to combined horizontal and vertical ground motions as in Cases (c) and (d) of Model II, the first five modes may be used for an adequate design.

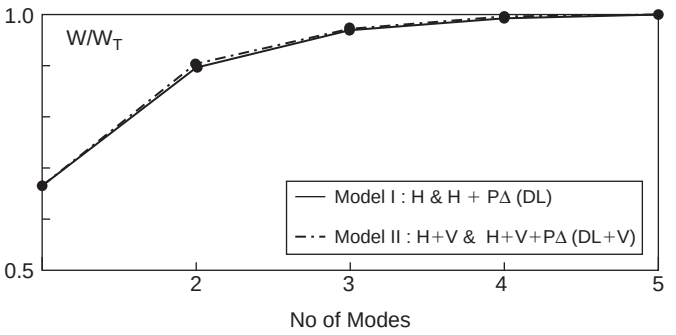


Figure 12 Ratio of kinetic and strain energies.

5.2 *Recursive optimality criterion techniques*

In this sub-section, efforts are emphasized on several items: the improvement of achieving optimal solution by using recursive procedures; the development of cost function including construction and damages; the assessment of various building code requirements through numerical results; and comparison of optimal cost functions. The recursive procedures have three levels which are described as follows (Cheng and Juang, 1988).

5.2.1 *First – level recursion*

This procedure is based on the envelope optimization technique described in section 5.1

5.2.2 *Second – level recursion*

The design results from the first recursion can be further improved by using the second level of recursion based on the constraint gradients for which the algorithm is of the following form:

$$(x_i)^{(v+1)} = (x_i)^{(v)} + s(\Delta x_i) \tag{100}$$

where s = the step size determining the rate of approach to the algorithm; and Δx_i = the changes of the members that can reduce the objective. Let us reduce the size of all of the members in a structure by a certain percentage. This change will reduce the structural stiffness, and, consequently, the active behavior constraint, g_j, is increased beyond the constrained surface by Δg_j. In order to bring it back to the constrained surface, one can increase the size of a certain number of members, say q. The magnitude of the increment of member size, Δx_i, is determined on the basis of the following assumptions: (1) the change in structural behavior is directly proportional to the change of member size; (2) the change of member size is inversely proportional to the member length because an increase in the size of a member with a larger length causes a larger increase in weight; and (3) the influence on the jth behavior constraint due to the unit change of

the ith member size, dg_{ji}, is directly proportional to the member size x_i. The resulting form is

$$\Delta x_i = \frac{\Delta g_j}{\sum_{q=1}^{m} R_q}\left(\frac{dg_{ji}}{l_i}\right); \quad R_q = \frac{1}{x_q l_q}(dg_{jq})^2 \tag{101a,b}$$

For multiple active behavior constraints, one should determine the required change in each element size separately for each constraint. The largest value of Δx_i shall be used for the actual change in the size of the ith element.

5.2.3 *Third – level recursion*

For multiple active constraints of displacements, stresses, and frequencies, it is necessary to find the Lagrange multipliers corresponding to the active constraints of the current design. In the third level of recursion procedures, it is to find the minimum value of λ for all active constraints of member *i*; the value is the then adopted in the following recursion relation:

$$(x_i)^{(k+1)} = (x_i)^{(k)}\left\{\frac{\max\left(\frac{\partial g_j(x)}{\partial x_i}\right)\min\left(\lambda_{ij}\right)}{\frac{\partial W(x)}{\partial x_i}}\right\}^{1/2} \tag{102}$$

in which min (λ_{ij}) is determined from all active constraints for an individual member *i* from which the maximum value of $\partial g_j(x)/\partial x_i$ is also obtained. Thus, the maximum and minimum search will yield the upper bound of the member size requirement for all active constraints for all conditions.

5.2.4 *Minimum cost*

For aircraft and automobile industries, the structural weight may not be only convenient but also essential in structural optimization. In seismic structural engineering, the construction and repair cost may be necessary to identify an optimum solution. Therefore a new objective including the costs of the structural members, paintings, connections, and damage is presented (Cheng and Juang, 1988).

A. STRUCTURAL MEMBER COST

The structural member cost may include the basic cost of common sizes used in constructions and the extra charge for the smaller and larger sizes of the members not commonly used. The cost function may be expressed as

$$C_{BE} = C_s\sum_{i=1}^{n}\rho_i A_i l_i + \sum_{i=1}^{n} 0.00916\rho_i A_i^{0.79} l_i \tag{103}$$

in which C_s = the unit price of steel ($/lb). The first term on the right side of Eq. (103) represents the basic cost due to the weight of each member and the second term is associated with the extra charge for any member having unusual size. Based on statistical

data, the extra charge is modeled in a function of the cross-sectional area as shown in the equation.

B. PAINTING COST

The amount of painting is measured according to the surface area of the members. Let C_{pt} be the cost of unit surface area and C_p the total painting cost of the structure, then

$$C_p = C_{pt}\left[\sum_{i=1}^{n_{bm}} \eta_i^{-0.54328}\left(\frac{8.5043}{2.09 - 1.212\eta_i^{2.17391} I_i} + 4.0712\right) l_i + \sum_{i=1}^{n_c} \eta_i^{-0.4808}\left(\frac{10.3312}{2.09 - 0.5565\eta_i^{1.9231} I_i} + 4.946\right) l_i\right] \tag{104}$$

in which n_{bm} and n_c, represent the number of beams and columns, respectively

C. CONNECTION COST

The connection cost is primarily concerned with welded plate beam-column connections, which include the costs of steel plate and welding.

$$(C_{pc})_i = 0.8796 C_{pl}\rho\eta_i^{1.087} I_i^2\left(\frac{10,800,000}{k_i} + 0.73301\eta_i^{1.087}\right) \tag{105}$$

in which C_{pl} = the unit price (\$/in^3) of the steel plate. The value of k_i is the initial slope of the moment-rotation curve of a connection for which the rotation is the relative angle between the elastic lines of the connected members at their point of intersection. The welding cost of beam i, $(C_{wc})_i$, may be expressed as the unit price of welding C_w times the volume of the welding materials:

$$(C_{wc})_i = C_w\rho(0.8919\eta_i^{1.087} I_i + 3.881 \times 10^{-3}\eta_i^{2.174} I_i^2 + 0.08643\eta_i^{1.6304} I_i^{1.5}) \tag{106}$$

Thus the connection cost for m_{bm} beams is given by

$$C_{con} = \sum_{i=1}^{n_{bm}} [(C_{pc})_i + (C_{wc})_i] \tag{107}$$

D. DAMAGE COST

The damage cost is limited to the costs of repairing nonstructural components including partitions and glasses. Thus the damage is expressed in structural drift at each floor and the damage repair costs on the ith floor is

$$(C_D)_i = 8.52 N_l (C_{nc})_i c_i \gamma \pounds \tag{108}$$

in which

$$\pounds = -\frac{1}{\vartheta}\left[a_{max}\exp(-\vartheta a_{max}) + \frac{1}{\vartheta}\exp(-\vartheta a_{max}) - \frac{1}{\vartheta}\right] \tag{109}$$

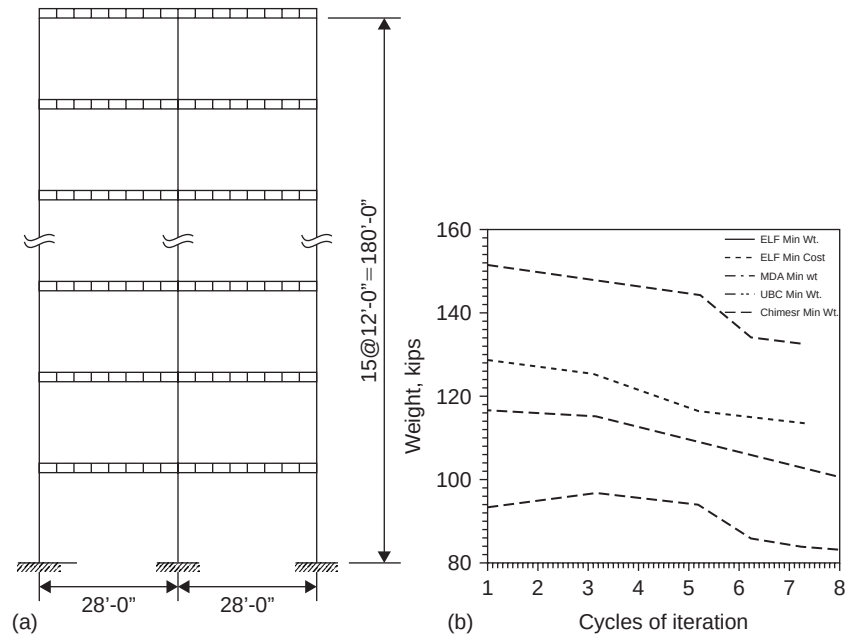


Figure 13 Example 9 – (a) 15-story 2-bay frame, (b) Five design cases.

where a_{max} = the maximum expected normalized ground acceleration in the lifetime of the structure; N_l is the life-time of the structure; $(C_{nc})_i$ is the construction costs of the damaged items on the ith floor; c_i signifies the structure's designed ground acceleration, a^*, story drift, Δ_i^*, as $c_i = \Delta_i^* / a^*$; γ and ϑ reflect seismic shocks within the specified range of ground acceleration as

$$n_0 = \gamma e^{-\vartheta a} \tag{110}$$

where n_0 represents earthquake frequency, a is the ground acceleration normalized by gravity acceleration, and ϑ signifies earthquake magnitude and seismic severity.

EXAMPLE 9 FIFTEEN STORY, TWO-BAY, UNBRACED FRAME FOR COMPARISON OF OPTIMAL DESIGN RESULT BASED ON BUILDING CODE REQUIREMENTS

The fifteen-story, two-bay structure shown in Fig. 13(a) was designed for ATC-03, UBC, and TJ-11-78. For ATC-03, the equivalent lateral force (ELF) procedures and modal analysis (MDA) procedures were used; furthermore, both the minimum weight and minimum cost were employed in the ELF design. According to ATC-03 requirements, the following parameters were used: a response modification factor of $R = 8$, a deflection amplification factor of $C_d = 5.5$, and an allowable story drift of $\Delta_a = 0.015 h_{sx}/C_d$ where h_{sx} is the story height below level x. The nonstructural dead load on each floor level was set at 100,000 lb. The AISC wide-flange sections and the three levels of recursions were used for the displacement constraint design.

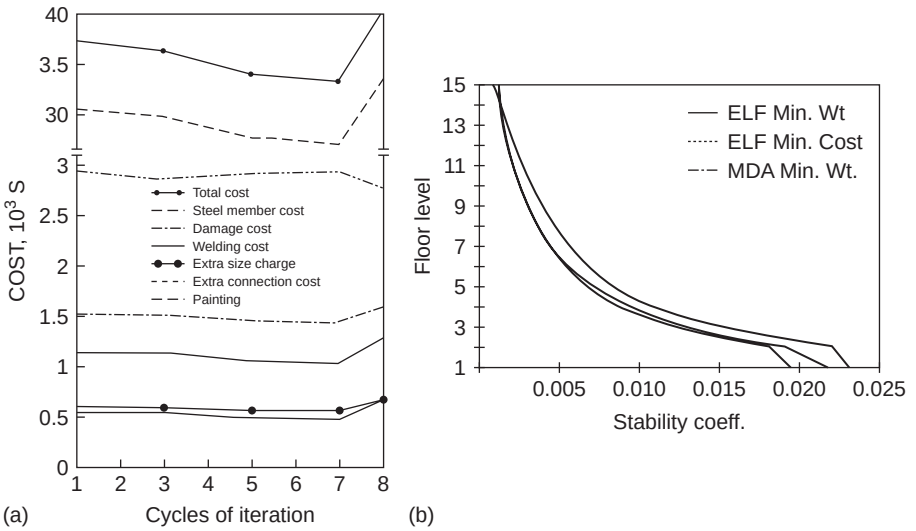


Figure 14 Example 9 – (a) Itemized optimal costs, (b) Comparison of stability coefficients.

For minimum cost, the unit prices of the steel, painting, steel at the connections, and welding metal are 0.24 \$/lb, 0.7986×10^{-3} \$/in^2, 0.3 \$/lb, and 5.5 \$/lb, respectively. The design was based on a ground acceleration of 0.4 g, and the maximum expected intensity of seismic excitations for 50 years, the lifetime of the structure, was assumed to be 0.5 g. the nonstructural cost is assumed to be 10% of the construction cost. In addition, rigid connections were assumed, therefore a large value of 9×10^9 was used as the initial stiffness of the connection, k.

The optimum weights of the five design cases are shown in Fig. 13(b). It is apparent that the UBC requires much less design than ATC-03 for which ELF, however, requires more weight than MDA, and the optimum solution of minimum weight design is less than the weight associated with the minimum cost case. The individual costs of the minimum cost design are shown in Fig. 14(a) from which one may observe that the design is primarily dominated by steel member costs, and that when the damage cost decreases the construction cost increases as shown in Cycles 7 and 8. The stability coefficients, θ, at the optimum solution are plotted in Fig. 14(b) for the design cases associated with ATC-03. The largest θ is 0.023, corresponding to the MDA minimum weight design which is much less than the critical value of 0.1 specified in ATC-03 (Cheng, 1986; Cheng and Juang, 1988).

5.3 *Generalized optimality criterion technique*

This technique is also based on Kuhn–Tucker necessary condition expressed as

$$\frac{\partial L}{\partial x_i} = \frac{\partial W(x)}{\partial x_i} + \sum_{j=1}^{l} \lambda_j \frac{\partial g_j(x)}{\partial x_i} = 0 \quad i = 1, \ldots, n \tag{111a}$$

with

$$\left.\begin{aligned}\lambda_j &\geq 0 \quad j = 1, \ldots, m\\ \lambda_j g_j &= 0 \quad j = 1, \ldots, m\end{aligned}\right\} \tag{111b}$$

Rearranging Eq. (111a) gives

$$T_i = -\sum_{j=1}^{l} \lambda_j \frac{\left(\frac{\partial g_j(x)}{\partial x_i}\right)}{\left(\frac{\partial F(x)}{\partial x_i}\right)} = 1 \quad i = 1, \ldots, n \tag{112a}$$

Linear recurrence relationship can be divided between cycle $(k+1)$ and k as

$$x_i^{(k+1)} = x_i^{(k)} \left(1 + \frac{1}{r}(T_i - 1)\right) \quad i = 1, \ldots, n_1 \tag{112b}$$

where n_1 = number of active design variables; $(n - n_1)$ = number of passive design variables; and r is the convergence control parameter. The approximate change in constraint can be expressed as

$$\Delta g_j = g_j(x + \Delta x) - g_j(x) = \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i}(\Delta x_i) \quad j = 1, \ldots, m \tag{113a}$$

Assume g_j is active, the change of Δx should force $g_j(x + \Delta x)$ to be zero, then combining Eqs. (112b) and (113a) yields

$$-g_j(x) = \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \frac{1}{r}(T_i - 1)(x_i^{(k)}) \quad j = 1, \ldots, m \tag{113b}$$

Note that the upper and lower limits of design variable affect the recurrence, thus (113b) should be modified to consider that effect as

$$-g_j(x) = \sum_{i=1}^{n_1} \frac{\partial g_j}{\partial x_i}(\Delta x_i) + \sum_{i=n_1+1}^{n_1} \frac{\partial g_j}{\partial x_i}(x_i^P - x_i^{(k)}) \quad j = 1, \ldots, m \tag{113c}$$

where x_i^P represents an element that becomes passive during kth iteration. In order to have $T_i = 1$, λ_j must also be active and the design variables x_i must be active for which the equation is shown as

$$\begin{aligned} &r g_j(x) - \sum_{i=1}^{n_1} \frac{\partial g_j(x)}{\partial x_i} x_i^{(k)} + r \sum_{i=n_1+1}^{n} \frac{\partial g_j(x)}{\partial x_i}(x_i^P - x_i^{(k)})\\ &= \sum_{t=1}^{m} \lambda_t \left(\sum_{i=1}^{n} \left(\frac{\frac{\partial g_j(x)}{\partial \delta_i} \frac{\partial g_t(x)}{\partial \delta_i}}{\frac{\partial W(x)}{\partial \delta_i}} \right) x_i^{(k)} \right) \quad j = 1, \ldots, m \end{aligned} \tag{114}$$

Eq. (114) is a general equation that can be used to find the Lagrange multipliers associated with the m active constraints. In order for these Lagrange multipliers to be valid, λ must be non-zero and positive value. For a negative Lagrange multiplier, the associated constraint should removed from the active set according to the Kuhn-Tucker conditions. In fact the negative Lagrange multiplier changes the direction of the constraint gradient vector indicating that enforcing this constraint will move the solution to a higher value of the objective. Note that scaling is an important procedure in optimality criterion method which is used to adjust the design variables such that some constraints will be active after scaling. Generally the scaling is the maximum ratio of current calculated constraints to the allowable lower or upper limit of constraints. It is therefore a scaling cycle can immediately bring the design to the constraint surface without tedious search procedures used in mathematical programming.

Note that the primary design variable is moment of inertia of a member about major axis, I_x, and that the structural weight is in function of member's cross section area, A. Thus, the secondary design variables must be expressed in terms of I_x. Based on regression numerical procedure, the equations are approximately determined from the AISC manual for wide-flange shapes for three groups: a) $I_x < 1500\,\text{in}^4$, b) $1500 \leq \text{I}_x < 12{,}100\,\text{in}^4$, and c) $I_x \geq 12{,}100\,\text{in}^4$. For example, the equations of group a) are

$$\left.\begin{aligned} I_y &= 0.0389 I_x^{0.925}; \quad J = 0.0221 I_x^{0.958} \\ A &= 0.5008 I_x^{0.487}; \quad S_x = 0.4531 I_x^{0.774} \\ S_y &= 0.0423 I_x^{0.732} \end{aligned}\right\} \tag{115}$$

For the optimization technique presented here, the gradients are derived in the following sub-sections (Cheng, 1986; Truman and Cheng, 1990).

5.3.1 *Static response gradients*

The algorithm uses the virtual load technique based upon the premise that the static displacements and stresses can be written as a linear combination of the structural displacements. This can be written as

$$u_j = \{b\}_j^{\text{T}} \{U\} \tag{116}$$

where, $\{b\}_j^{\text{T}}$, is the appropriate vector to enforce this relationship; u_j is the jth global displacement or stress; and {U} is the vector of the system's global displacements (see Fig. 15). The component of the gradient for the jth static displacement or stress constraint may be derived as

$$\frac{\partial u_j}{\partial x_i} = -\{v\}_j^{\text{T}} \frac{\partial [K_{Ei}]}{\partial x_i} \{U\} \tag{117}$$

$$\{v\}_j^{\text{T}} = \left\{b\right\}_j^{\text{T}} [K]^{-1} \tag{118}$$

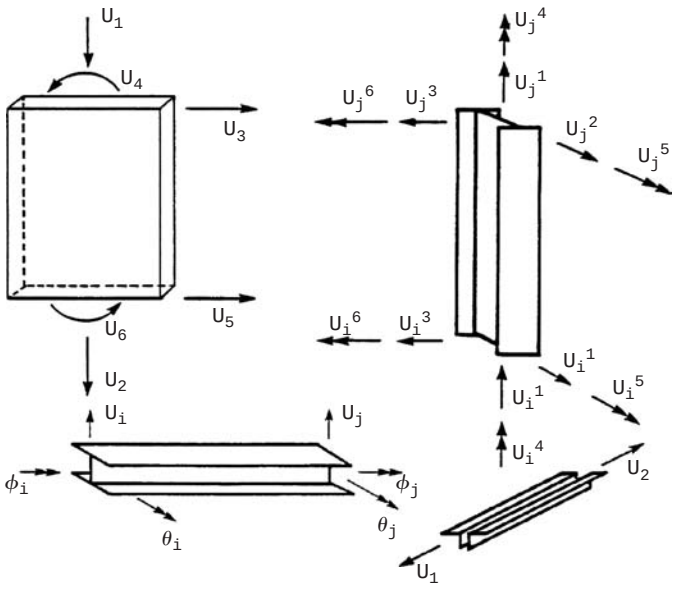


Figure 15 Member's and system's coordinates.

where $\{v\}_j$ is virtual displacement vector for jth constraint. The derivation of $\{b\}_j^{\mathrm{T}}$ depends on the nature of constraints and is given separately as follows

A. DISPLACEMENT

$$\{b\}_j^{\mathrm{T}} = [0\ldots,0,1,0,\ldots,0] \tag{119}$$

where there is a unit value at the jth location.

B. STRESSES

The stress constraints are element type dependent such as beams, columns, panels, etc. Herein, the beam-column stress constraint is illustrated for which the combined stress is

$$\sigma = \frac{P}{A} \pm \frac{M_x c}{I_x} \pm \frac{M_y d}{I_y} \tag{120}$$

where P is the axial load; M_x and M_y are the moments about the x and y axes, respectively; c and d are the appropriate distances from the neutral axes to the outer most fibers; and A, I_x, and I_y are the geometric properties. Eq. (120) can be expressed in terms of member-end displacements as

$$P = \frac{EA}{L}(U_i^1 - U_j^1) \tag{121}$$

$$M_{xi} = \frac{6EI_x}{L^2}(U_i^2 - U_j^2) + \frac{4EI_x}{L}\left(U_i^6 + \frac{1}{2}U_j^6\right) \tag{122}$$

$$M_{yi} = \frac{6EI_y}{L^2}(U_i^3 - U_j^3) + \frac{4EI_y}{L}(U_i^5 + \frac{1}{2}U_j^5) \tag{123}$$

where E is Young's modulus and L is the length of the beam-column. Substituting Eqs. (121) through (123) into Eq. (120) yields $\{b\}_j$ for the stress at end i of the member as

$$\begin{aligned}\{b\}_j^{\mathrm{T}} = &\left[\frac{E}{L} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad -\frac{E}{L} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\right] \\ &\pm \frac{Ec}{L^2}[0 \quad 6 \quad 0 \quad 0 \quad 0 \quad 4L \quad 0 \quad -6 \quad 0 \quad 0 \quad 0 \quad 2L] \\ &\pm \frac{Ec}{L^2}[0 \quad 0 \quad -6 \quad 0 \quad 4L \quad 0 \quad 0 \quad 0 \quad 6 \quad 0 \quad 2L \quad 0]\end{aligned} \tag{124}$$

where the signs are chosen according to the active stress found by Eq. (120) and j represents jth active constraint. The stress gradients of other types of elements can be similarly derived.

5.3.2 *Dynamic response gradients*

A. FREQUENCY CONSTRAINT GRADIENTS

Frequency constraint gradients are found by a direct differentiation of the free-vibration equation used to find the natural frequencies ω_j as

$$\frac{\partial \omega_j^2}{\partial x_i} = \{\Phi\}_j^{\mathrm{T}}\left[\frac{\partial [K]}{\partial x_i} - \omega_j^2 \frac{\partial [M]}{\partial x_i}\right]\{\Phi\}_j \tag{125}$$

where $\{\Phi\}_j$ is the jth eigenvector. The dynamic gradients are derived by direct differentiation of the generalized displacement equation as follows.

A. DYNAMIC DISPLACEMENT

For modal analysis, the dynamic displacement is resulting by a direct linear superposition of individual modes. The square root of the sum of the squares is used in this algorithm and can be written as

$$U_j = \left[\sum_{k=1}^{t} Y_k^2\right]^{1/2} \tag{126}$$

where k represents the kth eigenvector; Y_k represents the kth modal component of U_j; and t is the total number of eigenvectors used in the modal analysis. Using the chain rule, the gradient can be written as

$$\frac{\partial U_j}{\partial x_i} = \sum_{l=1}^{t} \frac{\partial U_j}{\partial Y_l}\frac{\partial Y_l}{\partial x_i} \tag{127}$$

Eq. (127) involves more work on sensitivity analyses of response spectrum (if response spectrum is used) as well as sensitivity analyses of eigenvalues and eigenvectors, which are omitted here.

B. DYNAMIC STRESS

The dynamic stress gradients are determined by using Eq. (116) for which

$$\frac{\partial u_j}{\partial x_i} = \{b\}_j^{\mathrm{T}} \frac{\partial \{U\}}{\partial x_i} \tag{128}$$

This calculation of $\partial\{U\}/\partial x_i$ is similar to that shown in Eq. (127).

5.3.3 *Drift gradients*

The drift is the relative displacement between adjacent structural floors. Thus the drift between floor levels m and n can be written in terms of the displacements as

$$\Delta_{mn} = \{b\}_{mn}^{\mathrm{T}}\{U\} = [0 \ldots, 0, 1, 0 \ldots, 0, -1, 0, \ldots, 0]\{U\} \tag{129}$$

where the positive 1 is in the nth location and the negative 1 is in the mth location. The virtual load approach would then find the component of the gradient of the drift as

$$\frac{\partial \Delta_{mn}}{\partial x_i} = -\{v\}_{mn}^{\mathrm{T}} \frac{\partial [K]}{\partial x_i}\{U\} \tag{130}$$

where

$$\{v\}_{mn} = [K]^{-1}\{b\}_{mn} \tag{131}$$

EXAMPLE 10 FIVE STORY IRREGULAR STRUCTURE SUBJECT TO FREQUENCY CONSTRAINTS

A five story L-shaped structure shown in Fig. 16(a) is used to illustrate for frequency (period) constraints. Each level has a translational mass of 0.31 k-s^2/in and a rotational

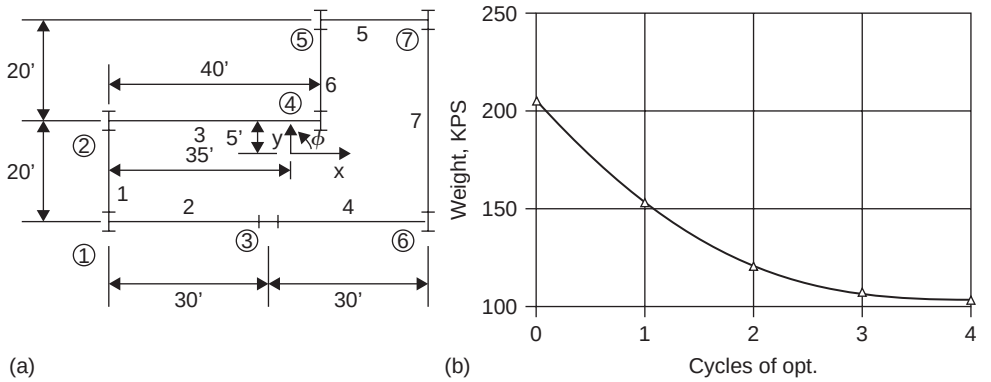


Figure 16 Example 11 – (a) 3D L-shaped building (b) Weight vs. optimization cycles.

mass of 16,403 k-s^2-in. the constraints consist of keeping the first period between the values of 0.75 and 1.0 sec., the second period below 0.50 sec. and the third period below 0.40 sec. while keeping each element size between 10.0 in^4 and 20,000.0 in^4.

The optimization was terminated due to less than a 5% change in weight between cycles. The final weight was 104 kips which is nearly a 50% reduction of the initial weight as shown in Fig. 16(b). All three of the period constraints became active with $T_1 = 1.02$ sec., $T_2 = 0.50$ sec., and $T_3 = 0.41$ sec., as shown in Fig. 17(a), while none of the side constraints became active. Note that the modes are coupled since the rigidity center and mass center do not coincide. The mode shapes are shown in Figs. 17(b) to (d). It is worthwhile mentioning that the ability to maintain certain frequencies or periods offers the advantages that structures can be designed in order to have specific frequencies which are in the favorable regions of the response spectrums.

EXAMPLE 11 TEN-STORY IRREGULAR STRUCTURE FOR SEISMIC RESPONSE PARAMETER STUDIES

The ten story setback structure shown in Fig. 18 is used for several different parameter studies based on ATC-03. ATC-03 recommends a number of important seismic building provisions including the seismic map areas for A_a and A_v which range from zero to seven corresponding to different areas within United States. A_a represents effective

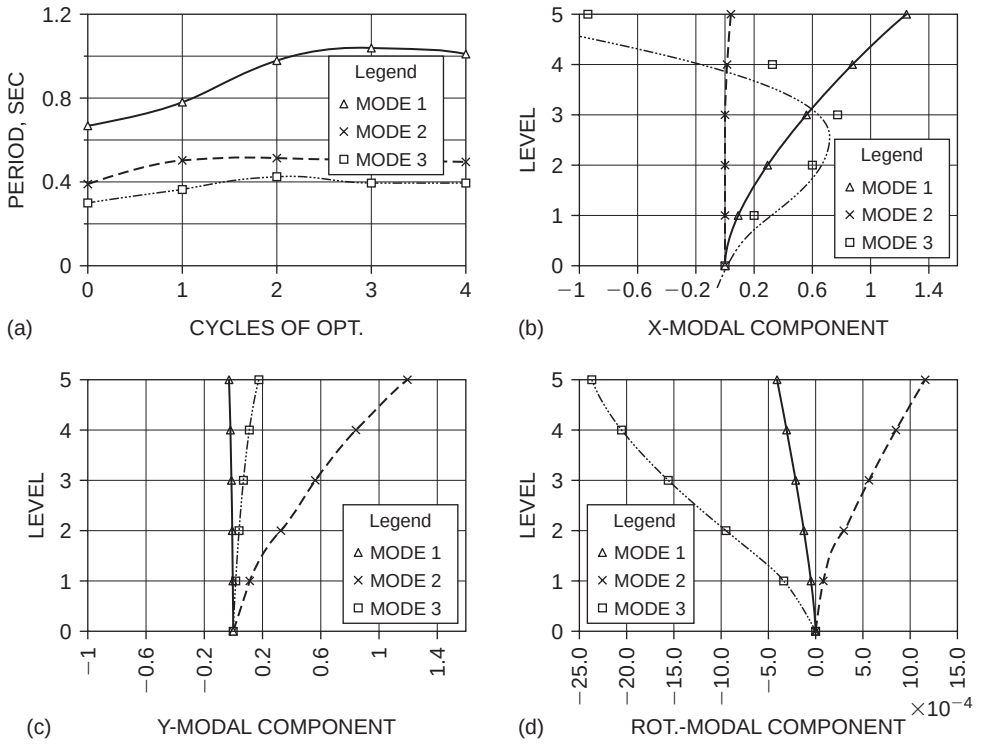


Figure 17 Example 11 – (a) periods, (b) X – component (c) Y – component (d) rotational component.

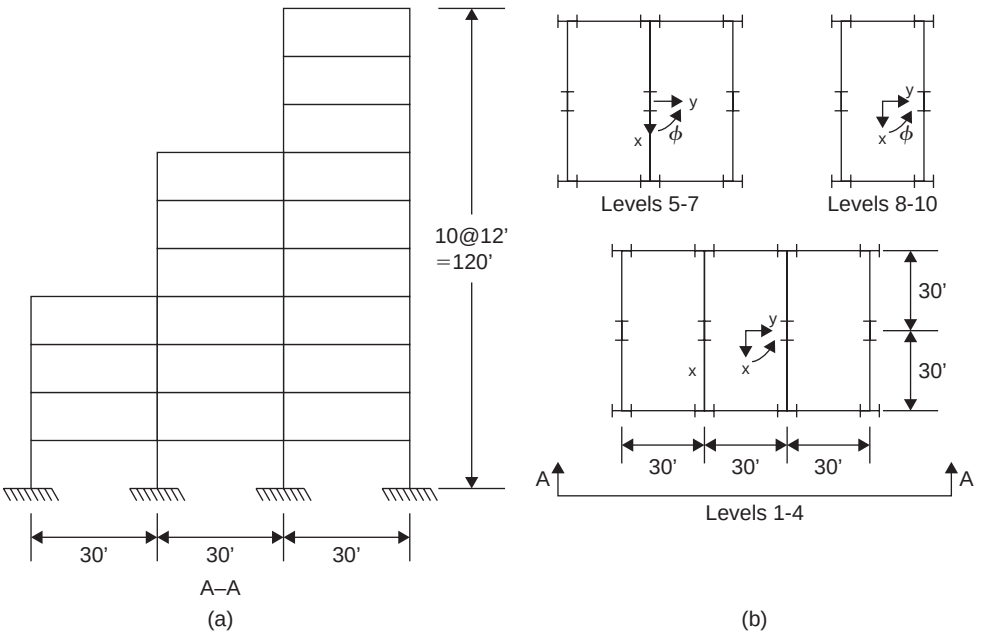


Figure 18 Example 11 – 10 story setback building (a) Elevation, (b) floor plan.

peak acceleration; Av signifies the effective peak velocity-related acceleration. Therefore, A_a and A_v were incorporated to take into account the resonance and distance effects associated with the seismic activity for any given structure. Area seven is the worst situation with respect to seismic activity (this corresponds roughly to a maximum ground acceleration of 0.4 g). A_v is developed to reflect that structures farther from the epicenter of the earthquake could very well be more sensitive to velocity due to the fact that ground motions tend to have an increase in duration and become more periodic with this distance. There are three soil types: Soil type 1, rock or stiff soil conditions where the soil depth is less than 200 feet; Soil type 2, deep cohesionless or stiff clay where the soil depth exceeds 200 feet; and Soil type 3, soft to medium-stiff clays and sand.

Seven combinations were considered with map areas ranging from 4 to 7. These seven combinations provide actual allowed combinations of these map areas. The optimal moments of inertia, I_x, of columns are given in Fig. 19(a) where the legend with three digit numbers signifies the map area and the soil type. First four modes are used in the modal analysis procedure. The design results are further given in Table 3 where (*) signifies the map area numbers for A_a and A_v; (**) represents loading conditions: Load 1 refers to a positive 5% eccentricity and Load 2 refers to a negative 5% eccentricity in the y-direction; and (***) identifies active displacement constraints, for instance x_{10}-x_8 indicates the displacements in the x-direction at the 8th through 10th floors active.

Table 3 Optimum Solution of 10-Story Setback Structure.

A_a^*	A_v^*	*Init. Wt. (kip)*	*Final Wt. (kip)*	*Period (sec)*	*Active Constraints*	
					*Load 1***	*Load 2***
7	7	594.2	515.8	1.477	–	x_{10}-x_8^{***}
6	6	516.5	476.4	1.717	x_{10}-x_8	x_{10}-x_5
5	6	516.6	454.9	1.716	x_{10}-x_8, x_6-x_5	x_{10}-x_5
5	5	423.9	387.3	2.206	x_{10}-x_3	x_{10}-x_3
4	6	515.9	473.6	1.725	x_{10}-x_8, x_6-x_5	x_{10}-x_5
4	5	424.1	387.4	2.116	x_{10}-x_8, x_6-x_5	x_{10}-x_5
4	4	368.5	314.4	2.489	x_{10}-x_5	x_{10}-x_5

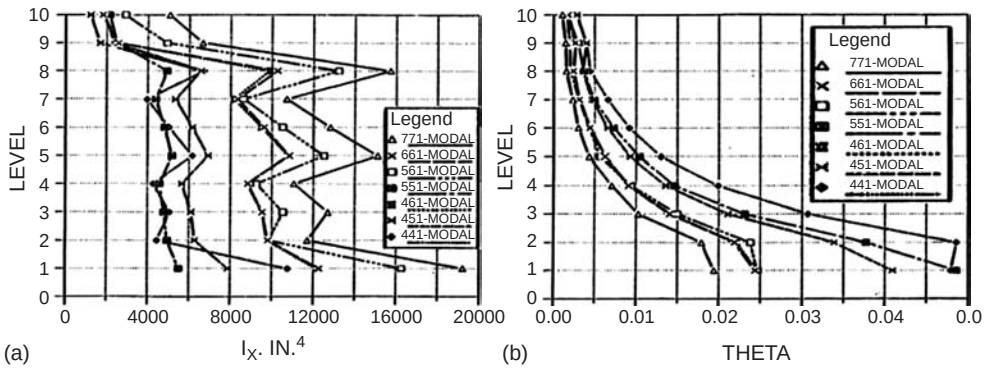


Figure 19 Example 11 – (a) Moment of inertia I_x, (b) Stability coefficients.

A. EFFECT OF VARIATION IN MAP AREAS FOR A_a AND A_v ON STABILITY FACTORS

This structure is designed by varying the values of A_a and A_v while holding the soil type constant (soli type 1) using modal analysis procedure. The seismic hazard exposure is assumed to be group 2, the response modification factor is 4.5, and the deflection amplification factor is 4.0. The translational and rotational masses for levels 8–10 are 0.839 k-s^2/in and 45,280 k-s^2-in; for levels 5–7 are 1.678 k-s^2/in and 144,895 k-s^2-in; for levels 1–4, 2.516 k-s^2/in and 353,180 k-s^2-in. In accordance with ATC-03, the gravity load should be considered as two loading cases in order to include the effects of the 5% eccentricity with respect to the major-axis loading. Thus structure's primary excitation in the x-direction at each level's mass center plus or minus the 5% eccentricity while the y-direction loading was considered to be thirty percent of the lateral loads determined for that direction. The dynamic displacement constraints are chosen as 0.45 in. per floor level. This was based on ATC-03 drift limit of 2.16 in. per floor, after inelastic effects are considered, or 2.16/4 = 0.54 in. for an elastic analysis (amplificator = 4).

The stability factors are shown in Fig. 19(b) for each of the seven cases and reveal that they are less than 0.05 which is well below the limit of 0.1. The larger stability factors occur within the lighter structures (lower effective peak acceleration values as expected.

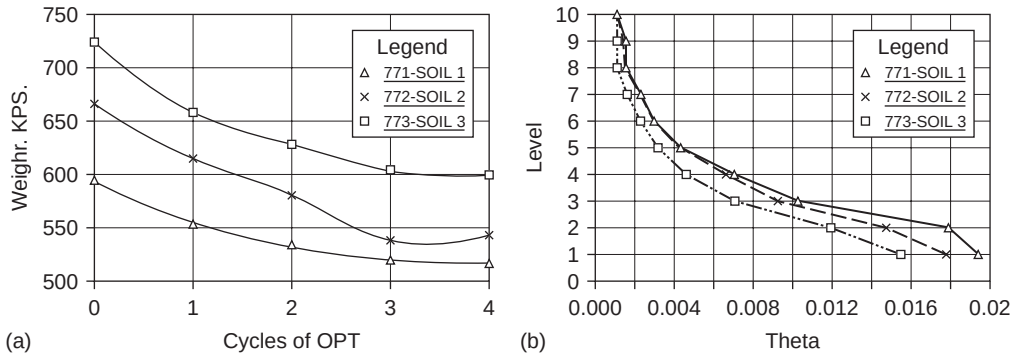


Figure 20 Example 11 – (a) Soil effects: weight vs. optimization cycles, (b) Stability coefficients.

B. EFFECT OF VARIATION OF SOIL TYPES ON STABILITY FACTORS

The ten-story setback structure is also designed with a variation in soil type of $S = 1,2,3$ for $A_a = 7$ and $A_v = 7$. The design results of optimal weight and stability function are given in Fig. 20(a) and 20(b), respectively. These results further provide the insight into the effect of different soil types on structural weight and particularly on stability factors which are much less than the limit recommended by ATC-03.

5.4 *Computer aided design of building optimization*

This is a computer based technique which is emphasized largely on structural analysis capacity embedded with optimal design features for tall buildings. The integral efforts of the technique are in the functionality of the computer program, OPTIMA, which has several practical features including: 1) similarity to two well-known commercial computer programs, ETABS and SAP2000; 2) centralized database for data management and operation; 3) graphical viewing tools for model checking; 4) data manipulation and modification; 5) display for visualizing the behavior and structural efficiency of buildings; 6) optimization techniques for practical building; and others. The optimization technique is relatively simple because the program is aimed at the design of tall buildings for which the control of lateral displacements (stiffness) is sufficient from practical point of view. Thus the objective function is expressed in terms of story drift. The algorithm is similar to the envelope optimality criterion technique with displacement constraint as presented in A of sub-section 4.1.2. As to the lateral wind force, the displacement constraints are not based on modal matrix presented in C of sub-section 4.1.2 based on Rayleigh quotient equation. Other constraints such as allowable stress and member sizes are considered to be secondary and treated member-by-members based on conventional design procedure after

system's drift limits are satisfied. For 3D steel framework, the floor displacement is expressed as

$$\delta = \sum_{i=1}^{N} \int_{0}^{L_i} \underbrace{\left(\frac{F_X f_X}{EA}\right)}_{\text{axial}} + \underbrace{\frac{F_Y f_Y}{EA_Y} + \frac{F_Z f_Z}{EA_Z}}_{\text{shear}} + \underbrace{\frac{M_Z m_Z}{GI_Z}}_{\text{torsion}} + \underbrace{\frac{M_Y m_Y}{GI_Y} + \frac{M_X m_X}{GI_X}}_{\text{bending}} \Big)_i \, dx \tag{132}$$

where $F_X, F_Y, F_Z, M_X, M_Y, M_Z$ = actual member forces and moments; f_X, f_Y, f_Z, m_X, m_Y, m_Z = virtual member forces and moments; $A, A_Y, A_Z, I_X, I_Y, I_Z =$ cross-sectional properties; E, G = material properties; X, Y, and Z are the cross sections axes.

As practical steel buildings require the use of commercial standard steel sections, the cross section properties A_Y, A_Z, I_X, I_Y, I_Z may be all expressed in terms of the axial area, A, which are derived as similarly presented in Eq. (115) as

$$\frac{1}{I_X} = \frac{C_{IX}}{A} + C'_{IX}; \quad \frac{1}{I_Y} = \frac{C_{IY}}{A} + C'_{IY} \tag{133}$$

where coefficients C and C' are determined by regression analysis.

For practical design, the steel members are limited to certain sizes such as W36, W30, W24, and W14, for which the relationship between cross-sectional area and strong moment of inertia, I_x, can be established. Thus the objective function of structural weight can be easily formulated. For reinforced concrete members, the deflection equation can be similarly obtained as shown in Eq. (132); and the objective function is based on cost of materials as well as construction as discussed in Eq. (36) for precast concrete members. Impressive design cases for various building systems are published (Chan, 1997).

EXAMPLE 12 DESIGN OF A TALL TRUSS FRAMEWORK

The 50-story 7-bay by 10-bay trussed framework with 5,400 members shown in Fig. 21(a) is designed with the following conditions: story height = 12 ft; typical bay width = 15 ft; material mass density = 0.49 kip/cu. ft; Young modulus, E = 29,000 ksi; Shear modulus, G = 11,200 ksi; columns are limited as W14 shapes (typical), 2W14 shapes (core), 1 column/2 stories; diagonals are limited as W24 shapes, 1 diagonal/2 stories; beams are limited as W_{24} shapes, and 1 type of beam/2 stories. Interstory drift limit is 1/400. The optimization results are shown in Fig. 21(b) where the optimal weight is 11,244 kips; and the selection of economic steel sections (discrete) close to optimum designed members yields 11,316 kips corresponding to 15.41 lbs. of steel per square feet of floor area.

6 Genetic algorithm and multiobjective optimization

6.1 *Brief introduction of genetic algorithm*

Genetic algorithms (GAs) (Goldberg 1989), based on a biological evolution mechanism and Darwin's survival-of-the-fittest theory, are intelligent search methods for solving complex problems. GAs belong to the class of stochastic search optimization methods, for which the decisions made in most computational steps are based on random number generation. The algorithms use only the function values in the search process to make

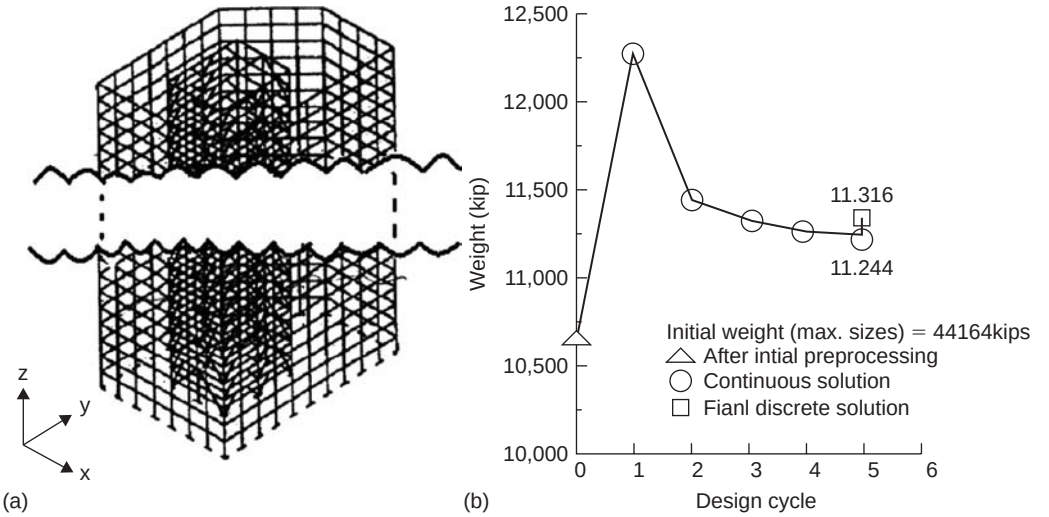


Figure 21 Example 12 – (a) Fifty-story framework (b) Optimal weight.

progress toward a solution without regard to how the functions are evaluated. Continuity or differentiability of the problem functions is not needed, thus the algorithms are very general and may be used to determine global optimum solutions. Because GAs are loosely parallel biological evolution and closely adapt the microbiologic terminology to mimic genetic operations, a brief explanation of the relationship between structural design languages and that of microbiology is relevant and given as follows: the term Population signifies a set of design points at the current iteration representing a group of designs as potential solution points; then iteration of the genetic algorithm is called a generation; a Chromosome represents a design of the system, whether feasible or infeasible; and the term Gene is used for a scalar component of the design vector; i.e., it represents the value of a particular design variable. Generally the procedure is to start with a set of designs, randomly generated allowable values for each design variable. Each design is also assigned a Fitness value using the cost function or penalty function depending on whether it is a constrained or non-constrained problem. From the current set of designs, a subset is selected randomly for better fit members of the set. Random processes are used to generate new designs using the subset of designs. Since more fit members of the set are used to create new designs, the successive sets of designs have a higher probability of having designs with better fitness values. The process is continued until a stopping criterion is met. More details are available form various publications.

6.2 *Brief introduction of multiobjective optimization and game theory*

Most real-world design optimization problems in structures are multimodal. There often exist several objectives to be considered by the designer. Usually these objectives are conflicting rather than complementary. Even though a few methods are available

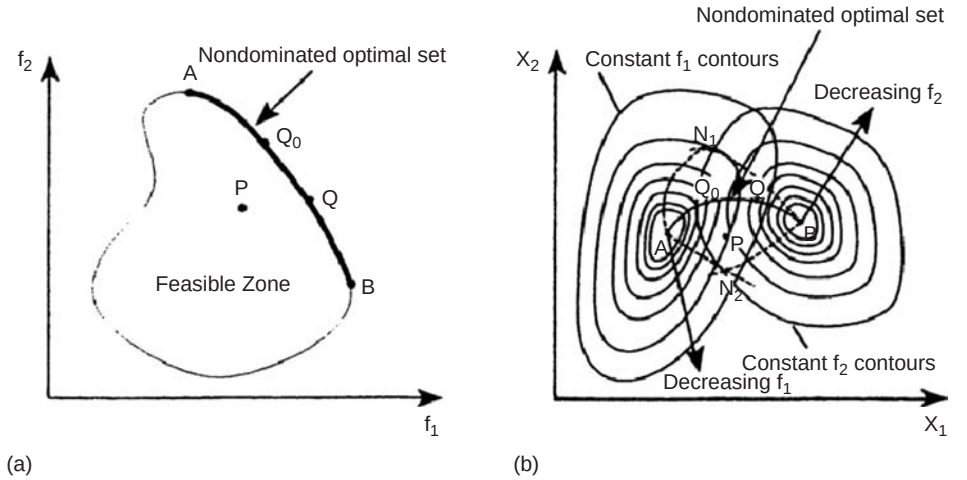


Figure 22 (a) Objective function space (b) Game solution.

for multiobjective optimization such as weighted min-max and weight global criterion, most of them often depend on function continuity. GAs is thus introduced here along with game theory. Game theory has been developed for both cooperative and non-cooperative games. In the latter each player of the game acts independently in an effort to maximize his own payoff, but the end result may not be favorable to the other. A cooperative game means that the players agree to form coalitions under the expectation that a mutually beneficial outcome can be obtained. The concept illustrated in Fig. 22. Observing Fig. 22(a) reveals that moving point Q_0 to Q increases the payoff of f_1; however, it decreases the payoff of f_2 at the same time. Point P is not a nondominated point; therefore, P moves to Q_0 which improves the payoffs of f_1 and f_2 simultaneously. Fig. 22(b) shows the two-objective optimum problem. The dotted lines passing through A and B stand for the loci of maximizing f_1 and f_2 for fixed values x_1 and x_2, respectively. Intersection points, N_1 and N_2, of the lines are candidates for the two-objective maximization problem assuming that the players, whose payoffs are f_1 and f_2, respectively, are in a non-cooperative game.

A multiobjective optimization problem (MOP) can be cast as a cooperative game problem in which it is assumed that one player is associated with an objective. The objective function f_i can be regarded as the payoff of the ith player. With cooperative multiobjective optimization, the "compromise solution" should make sure that each objective obtains its maximum possible value even though each objective cannot arrive at its own best value. Optimal tradeoff among the objectives is sought by using the concept of game theory as follows (Cheng and Li, 1997). First, the m individual objective functions are minimized respectively subject to given constraints.

$$\left.\begin{array}{ll} \text{Minimize} & F_i(x) \\ \text{Subject to} & g(x) \leq 0 \end{array}\right\} \tag{134}$$

For each objective function f_i, an optimal solution x_i^*, is obtained; then a pay-off matrix is constructed. The best and worst values in the Pareto set can be obtained from the pay-off matrix as

$$\left.\begin{aligned} f_{i,\min} &= f_i(x_i^*) \qquad j = 1,\ldots,m \\ f_{i,\max} &= Max\left[f_1(x_j^*)\right] \quad j = 1,\ldots,m \quad i = 1,\ldots,m \end{aligned}\right\} \tag{135}$$

Consequently the ith objective function should not expect a value better than $f_{i,min}$, but not worse than $f_{i,\max}$. Thus a substitute objective function can be constructed as

$$S = \prod_{i=1}^{m} \frac{[f_{i,\max} - f_1(x)]}{[f_{i,\max} - f_{i,\min}]} = \prod_{i=1}^{m} \bar{f}_i(x) \tag{136}$$

Maximizing the function S produces a solution that results in optimal compliance with multiple objectives subject to given constraints. The solution is a Pareto optimum and stands for the "rational compromise" of the conflicting objectives (Cheng and Li, 1997; Cheng, Ang, Lee, and Li, 1999).

EXAMPLE 13 DESIGN OF THREE-STORY FRAME FOR SINGLE AND MULTIPLE OBJECTIVE FUNCTIONS

The 3-story frame shown in Fig.. 23(a) is designed for El Centro, N-S, May 18, 1940 earthquake; live and dead load of 13 kip/ft at each story, and lateral forces of $F_1 = 8$ kip, $F_2 = 12$ kip, and $F_3 = 8$ kip. Weight density and elastic modulus of steel are 487 lb/ft^3 and 29000 ksi, respectively. The objective functions are: 1) minimum weight; 2) minimum input energy; 3) minimum weight and strain energy; 4) minimum weight and

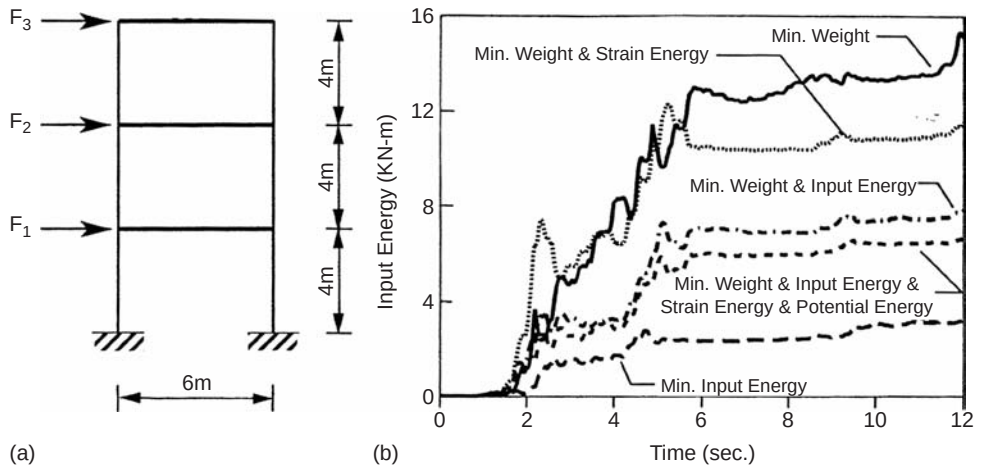


Figure 23 Example 13 – (a) 3-story frame, (b) Single and multiobjective functions.

seismic input energy; and 5) minimum weight, strain energy, potential energy, and input energy. The constraints are

$$\left.\begin{aligned} &|\sigma_i| \le 24\,\text{ksi}; |\delta_i| \le 1/400\,h_i; \quad T_1 \ge 0.3\,\text{s}; \\ &24 \le I_i \le 4805\,\text{in}^4; \quad 0.5 \le k_{i+1}/k_i \le 1 \end{aligned}\right\} \tag{137}$$

where σ_i, δ_i and k_i are column stress, relative story displacement and story, respectively, and $i = 1, 2, 3$. T1 is the fundamental natural period of the structure. After optimum design, the structural response is then studied for the same earthquake. A comparison response of top-story displacements is shown in Fig. 23(b) which reveals that when the structure is optimized for minimum weight the dynamic response is very strong as reflected from the input energy. However, the design resulting from optimization of earthquake input energy only yields much smaller dynamic response due to demanding of heavier structural weight. Thus multiobjective optimization can achieve a compromise solution from conflicting objectives.

EXAMPLE 14 OPTIMUM DESIGN FOR SELECTION OF A STRUCTURAL SYSTEM

For a ten-story building shown in Fig. 24(a), MOP is applied to evaluate three choices: steel frame, reinforced concrete frame, and composite system of reinforced concrete (1st – 4th floor) and steel (5th – 10th floor). The structural weight and construction costs are available in professional community (Means, 1997). Design variables are defined as follows: At each story, cross-sectional areas of all the beams are identical. Cross-sectional areas of the internal columns are the same, and the two outside columns are identical. Structural members and member locations of first and second stories are the same as those of the third and fourth, fifth through seventh, and eighth through tenth stories.

For seismic design, calculation of the earthquake force is based on UBC requirements (UBC, 1994), and the following assumptions are made. The building is located in seismic zone 2B with factor $Z = 0.20$; site coefficient $S = 1.2$, and importance factor $I = 1.0$. The building is a moment-resisting frame type with factor $R_w = 6$ for the steel frame, and $R_w = 5$ for the RC and the composite frame. The building is in the occupancy category of seismic zone 2, static lateral force procedure in UBC can be used.

Assume that the total superimposed load is 40 psf for the roof and 75 psf for the floor. The reinforced concrete flat plate without drop panels is used for the roof and the floor, the total loads of the roof and the floor plus the superimposed loads are $q_r = 127$ psf and $q_f = 175$ psf, respectively. Material properties are as follows. Weight density of reinforced concrete is $\rho_c = 150$ pcf. Yield stress of steel column and beam is $F_y = 36$ ksi. Elastic modulus of steel is $E_s = 29{,}000$ ksi. Compressive strength of concrete is $f'_c = 4$ ksi. Elastic modulus of concrete is $E_c = 3{,}605$ ksi. Yield stress of reinforcement is $f_y = 60$ ksi.

Objective functions comprise construction cost and potential energy for steel, reinforced concrete, and composite frames. The construction cost is based on professional practice that: The metal framing cost is depending on number of frame story. For instance 7 to 15 story frame is \$1,704 per ton including A36 steel; high-strength bolts; fabrication and delivery; unloading and sorting; erection equipment; field erection

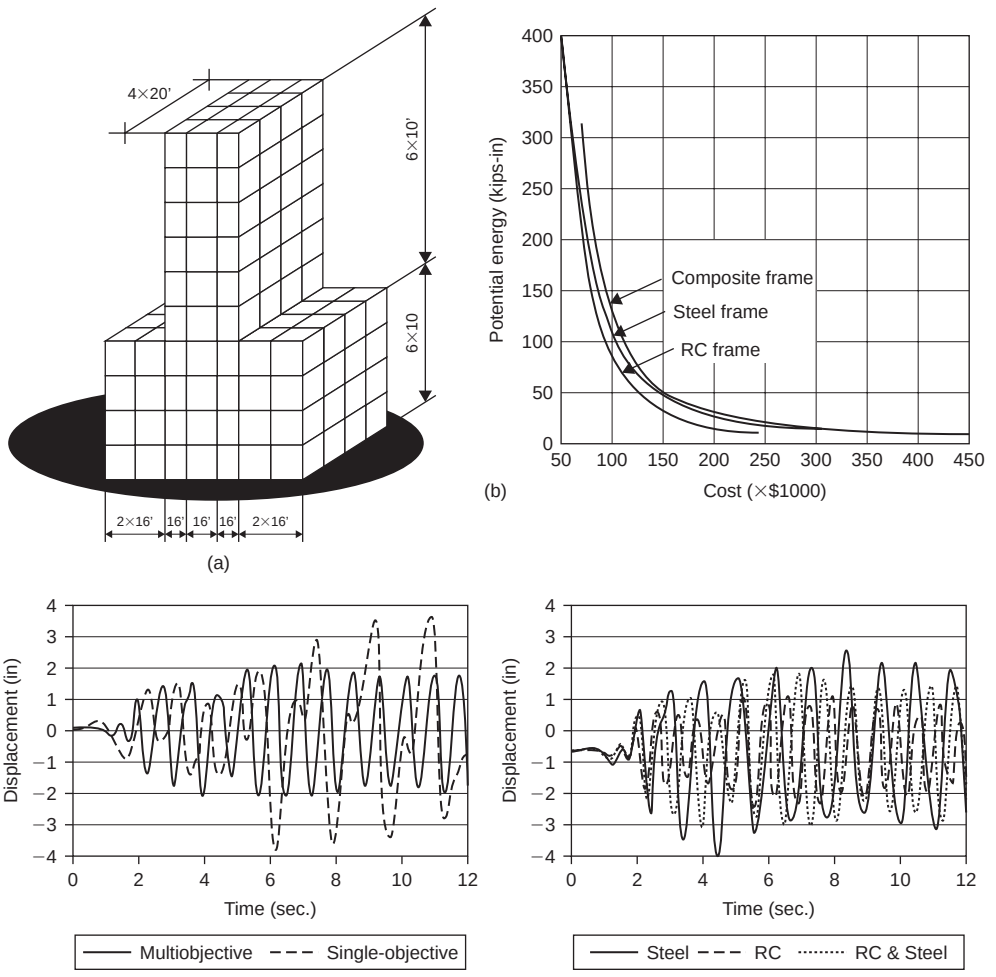


Figure 24 Example 14 – (a) Steel, R.C., composite building (b) comparison of optimal results (c) (d).

labor; and subcontractors overhead and profit. The cost parameters of the reinforced concrete members include material, mix, form, location, reinforcing, transportation, and subcontractors' overhead and profit are different for columns and beams. For instance, the cost is \$1775 and \$990 for columns of 12″ by 12″ and 16″ by 16″, respectively. As to beams the cost depends on span length and live load in kips per linear foot: for span length 16 ft the cost is \$457 and \$423 per cubic yard based on live load 2 to 3 and 4 to 5 kips per feet, respectively. The aforementioned cost is based on that the columns are designed with the building code suggested average reinforcement ratio and that the formwork is used four times. The constraints are allowable stress and allowable drift as well as upper and lower bound of member cross section.

Constraints are expressed in Eqs. (138) through (140) for steel, R.C. and composite systems, respectively

$$\left.\begin{aligned}&\left\{F_1=\sum_{i=1}^{m} f_{si};\quad F_2=\sum_{i=1}^{n} P_i\Delta_i\right\}\\&f_{bi}\le F_b,\quad i=1,\ldots,m_2;\quad \tfrac{f_a}{F_a}+\tfrac{f_b}{F_b}\le 1,\quad i=1,\ldots,m_1\\&\Delta_j\le\Delta_a,\quad j=1,\ldots,n;\quad A_i^L\le A_i\le A_i^u,\quad i=1,\ldots,m\end{aligned}\right\}\tag{138}$$

$$\left.\begin{aligned}&\left\{F_1=\sum_{i=1}^{m_1} f_{ci}+\sum_{i=1}^{m_2} f_{bi},\quad F_2=\sum_{i=1}^{n} P_i\Delta_i\right\}\\&\Delta_j\le\Delta_a,\ j=1,\ldots,n;\quad A_i^L\le A_i\le A_i^u,\quad i=1,\ldots,m\end{aligned}\right\}\tag{139}$$

$$\left.\begin{aligned}&\left\{F_1=\sum_{i=1}^{n_1} f_{si}+\sum_{i=1}^{n_2} f_{ci}+\sum_{i=1}^{n_3} f_{bi};\quad F_2=\sum_{i=1}^{n} P_i\Delta_i\right\}\\&f_{bi}\le F_b,\quad i=1,\ldots,m_4;\quad \tfrac{f_a}{F_a}+\tfrac{f_b}{F_b}\le 1,\quad i=1,\ldots,m_3\\&\Delta_j\le\Delta_a,\quad j=1,\ldots,n;\quad A_i^L\le A_i\le A_i^u,\quad i=1,\ldots,m\end{aligned}\right\}\tag{140}$$

where m = number of total members, m_1 = number of columns, m_2 = number of beams, n = number of structural stories. Δ_a = allowable drift (0.48 in.) F_a and F_b = allowable axial compressive and bending stresses, respectively. $A_i^L = 8\,\text{in}^2$ $A_i^u = 125\,\text{in}^2$.

The optimum design results are shown in Fig. 24(b), where the curves reveal a tradeoff between structural cost and structural systems.

Response performance of these three structures is studied. The structures are selected at the same cost of \$118,000; each of them undergoes time-history analysis with N-S El Centro May 18, 1940 earthquake. One of the essential response parameters is top-story displacement which is compared in Fig. 24(d). The RC frame has the smallest displacement response, and the steel frame has the largest.

The dynamic responses of single-objective optimum designs and multiobjective optimum designs are further compared for the composite system. The system is now designed with the same cost of \$118,000. The performance of both structures is studied under the same earthquake excitation. Fig. 24(c) shows that displacement resulting from multiobjective function is much smaller than that due to single objective function; maximum value at the top story decreases from 3.924 inches to 2.237 inches and maximum sway distance at the top story decreases from 7.384 inches to 4.412 inches.

6.3 *Genetic algorithm for multi-criteria cost optimization*

The main features of this approach include: A) formulation of cost functions; for instance the cost of steel structures are classified into three cases as 1) cost of the structure (cost rolled sections to account for different rolled shapes), 2) weight of steels, and 3) number of different section types to be used. Apparently these three functions are nebulous; a fuzzy function is used for each case with weighting coefficients of w_C, w_W and w_T assumed for cases 1, 2, 3, respectively. The sum of the weighting

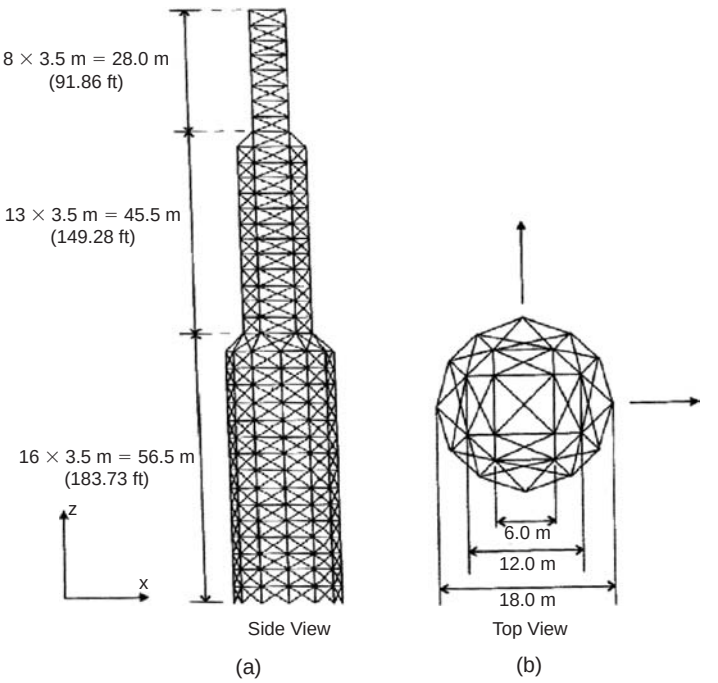


Figure 25 Example 15 1310-bar space truss.

coefficients is equal to 1. B) The members are designed using AISC specifications and therefore they are chosen from giving groups of available wide flange sections. Some sophisticated structures are designed by Adeli and his coworkers for static gravity loads and lateral forces. An illustration is given in Example 15.

Example 15 50-story Space Truss Design

The steel space truss shown in Fig. 25 of 1310 bars is designed for the following two criteria: Case A: $w_C = 0.01$; $w_W = 0.495$; $w_T = 0.495$ ($N_i = 5$ to 7; minimum material cost design) and Case B: $w_C = 0.495$; $w_W = 0.01$; $w_T = 0.495$ ($N_i = 5$ to 7; minimum material weight design) where N_i is the number of potential candidates for achieving the minimum optimal results. A higher number of N_i helps in choosing similar shapes used for other members because a wider range of candidates shapes facilitates the selection of already used shapes. The weighting coefficients are decided in accordance with designer's practical experience.

For columns and axially loaded members, A572 Grade 50; W6, W8, W10, W12, and W14 shapes are used. Similarly, for beams, W16, W18, W21, W24, W27, W30, W33, W36, and W40 shapes from the same grade of steel are used. The design constants used are modulus of elasticity E = 29,000 ksi, specific weight $\rho = 490.0$ lb/ft^3, and yield stress $F_y = 50$ ksi. The displacement constraint is the horizontal drift of ± 19.69 in (0.4%) of the structure's height. The basis of design is the AISC ASD code (Adeli and Kamal, 1986; Sarma and Adeli, 2000).

The minimum material cost solution of Case A for this example is not substantially less expensive than the minimum weight solution of Case B. This can be explained partly by the fact that, in this particular space truss example, the number of section types used in the minimum cost design is close to the number of section types used in the minimum weight design. Also, the section database included a relatively small number of wide flange shapes. Nevertheless, the example is only intended to illustrate the approach.

Type of study	*Material cost of the structure*	*Weight of the structure*	*Total Section types*
Case A for $N_i = 5,6,7$	$1,678,424	920.2 kips	17
Case A for $N_i = 5,6,7$	$1,688,853	918.1 kips	20

7 Summary

The paper introduces various algorithms and their essential formulations for structural optimization currently in vogue with extensive numerical examples to illustrate the advantages of using individual algorithms. The inclusion emphasizes on optimum design of building systems subjected to earthquake forces.

Acknowledgements

The paper is prepared with the support of the author's Curators' Professorship from the University of Missouri. The author regrets not being able to include a number of valuable publications because of the paper length constraint.

References

Adeli, H. and Kamal, O., Efficient Optimization of Space Trusses, *Computers and Structures,* 24(3), 501–511, 1986

AISC, *Manual of Steel Construction – Allowable Stress Design*, American Institute of Steel Construction, Chicago, Illinois, 1989

Ang, A.H-S., and Tang, W.H., *Probability Concepts in Engineering*, 2nd Edition, J. Wiley & Sons, Inc., New York, 2007

Applied Technology Council, *Tentative Provisions for the Development of Seismic Regulations for Buildings, ATC-3-06*, National Bureau of Standards, Washington D. C., 1978; also BSSC (1985)

Arora, J.S., *Introduction to Optimum Design*, 2nd Edition, El-Sevier Inc., California, 2004

Chan C.M., How to Optimize Tall Steel Building Frameworks, *ASCE Manuals and Report on Engineering Practice No. 90 – Guide to Structural Optimization*, ASCE, 165–196, 1997

Chang, C.C., Ger, J.R. and Cheng, F.Y., Reliability-Based Optimum Design for UBC and Non-deterministic Seismic Spectra, *Journal of Structural Engineering*, ASCE, Vol. 120, No. 1, pp. 139–160, 1994

Cheng, F.Y., *Matrix Analysis of Structural Dynamics – Applications and Earthquake Engineering*, Marcel Dekker, Inc., also CRC Press, Taylor and Francis Group, Boca Raton, Florida, 2001

Cheng, F.Y. (ed.), *Recent Developments in Structural Optimization*, ASCE, 1986

Cheng, F.Y., Venkayya, V.B., and Khachaturian, N., et al., *Computer Methods of Optimum Structural Design*, Vols. 1 and 2, University of Missouri-Rolla, Rolla, Missouri, 1976

Cheng, F.Y. and Botkin, M.E., Nonlinear Optimum Design of Dynamic Damped Frames, *J. of Structural Div.*, ASCE., Vol. 102, No. ST3, pp. 609–628, March, 1976

Cheng, F.Y., and Srifuengfung, D., Optimum Structural Design for Simultaneous Multicomponent Static and Dynamic Input, *International Journal for Numerical Methods in Engineering*, Vol. 13, No.2, pp. 353–372, 1978.

Cheng, F.Y., *Optimum Design of Reinforced Concrete and Steel Structures Subjected to Static and Seismic Loads*, National Taiwan Institute of Technology, Taiwan, Vols. I (242 pages) and II (375 pages), 1986

Cheng, F.Y. and Chang, C.C., *Safety-Based Optimum Design of Nondeterministic Structures Subjected to Various Types of Seismic Loads*, NSF Report, U.S. Department of Commerce, National Technical Information Service, Virginia, NTIS No. PB90-133489/AS (326 pages), 1988

Cheng, F.Y. and Juang, D.S., Assessment of Various Code Provisions Based on Optimum Design of Steel Structures, *Journal of Earthquake Engineering and Structural Dynamics*, Vol. 16, pp. 46–61, 1988

Cheng, F.Y. and Li, D., Multiobjective Optimization Design with Pareto Genetic Algorithm, *Journal of Structural Engineering*, ASCE, Vol. 123, No.9, pp. 1252–1261, 1997

Cheng, F.Y., Ang, A.H-S., Lee, J.R. and Li, D., Genetic Algorithm for Multiobjective Optimization and Life-Cycle Cost, *Structural Engineering in the 21st Century*, ASCE, pp. 484–489, 1999

Cheng, F.Y., Jiang, H.P. and Lou, K.Y., *SMART STRUCTURES – Innovative Systems for Seismic Response Control*, CRC Press, Taylor and Francis Group, Boca Raton, Florida, 2008

Ellingwood, B., and Ang, A.H-S., *A Probabilistic Study of Safety Criteria for Design*, Structural Res. Ser. No. 387, University of Illinois, 1972

Ellingwood. B. and Culver, C., Analysis of Live Loads in Office Buildings, *J. of Structural Div.*, ASCE, Vol. 103, No. ST8, pp. 1551–1560, 1977

Frangopal, D.M. and Cheng, F.Y. (Eds.), *Advances in Structural Optimization*, ASCE, 1996

Goldberg, D.E, *Genetic Algorithms is Search Optimization and Machine Learning*, Addison – Wesley, Reading, Massachusetts, 1989

Khot, N.S., Berke, L., and Venkayya, V.B., Comparison of Optimality Criteria Algorithms for Minimum Weight Design of Structures, *AIAA J.*, 17(2), 182–190, 1979

Means, *Building Construction Cost Data/RS Means* (Annual) RS Means Company, Kingston, Massachusetts, 1997

Mitchell, G. and Woodgate, R., *Floor Loadings in Office Buildings the Result of a Survey*, Current Paper 3/71, Building Research Station, Garston, Watford, England, 1971

Mohraz, B., Hall, W. and Newmark, N., *A Study of Vertical and Horizontal Earthquake Spectra*, Div. of Reactor Standards, U. S. Atomic Energy Commission Washington, D.C., 1972.

NBS 577, *Development of a Probabil ity Based Load Criterion for American National Standard A58*, National Bureau of Standards, 1980

Romstad, K.M. and Wang, C.K., Optimum Design of Framed Structures, Journal of Structural Division, ASCE, Vol. 94, No. ST12, 1968

Romstad, K.M., Hutchinson, J.R. and Runge, K.H., Design Parameter Variation and Structural Response, International Journal for Numerical Methods in Engineering, Vol. 5, Issue 3, pp. 337–349, 1972

Sarma, K.C. and Adeli, H., Fuzzy Genetic Algorithm for Optimization of Steel Structures, *Journal of Structural Engineering, ASCE*, 126(5), 596–604, 2000

Truman, K.Z. and Cheng, F.Y., Optimum Assessment of Irregular 3-D Seismic Buildings, *Journal of Structural Engineering*, ASCE, Vol.116, No.12, pp. 3324–3337, 1990

Venkayya, V.B., Khot, N.S., and Reddy, V.S., *Energy Distribution in an Optimum Structural Design*, Technical Report AFFDL-TR_68-156, Wright-Patterson Air Force Base, Ohio, 1969

Venkayya, V.B. and Cheng, F.Y., Resizing of Frames Subjected to Ground Motion, *Proc. of the International Symposium on Earthquake Structural Engineering*, University of Missouri-Rolla, Vol. 1, pp. 597–612, 1976

Chapter 5

Quality assurance and risk reduction in foundation engineering

Wilson H. Tang & Limin Zhang

Department of Civil Engineering, Hong Kong University of Science and Technology, Hong Kong

ABSTRACT: Foundation design and construction can be divided into three phases: site investigation, design, and construction. In each of the phases, engineers attempt to ensure the quality and reliability of the foundation. This paper outlines the important roles reliability theory and quality assurance play in each of these phases. In the site investigation phase, reliability theory is used to enhance the planning of investigation boreholes and in-situ tests so that any anomalies and key ground features can be revealed more effectively. In the design phase, evaluation of design analysis models and model parameters, as well as verification of new construction workmanship are of great concern. A recommended way to risk management is to use an integrated approach that combines site investigation, design, and construction. In this approach, an early test program, additional borehole boring, and a verification test program at the end of construction are implemented as means to manage various sources of uncertainty. The practices can be expressed in a Bayesian framework of updating information. In the construction phase, site supervision, monitoring, and quality assurance tests such as integrity tests, PDA tests, and static load tests are routinely conducted to ensure the quality and safety of foundations. It is possible to apply reliability theory to effectively utilize both quantitative (e.g. load tests, movements and ground water monitoring) and qualitative information (e.g. site supervision and integrity tests in the form of "pass" or "no anomaly" statements) for updating model errors and design parameters, and assisting in decision-making.

1 Introduction

Foundation design and construction can be divided into three phases: site investigation, design, and construction. In each of the phases, engineers attempt to ensure the quality and reliability of the foundation. Yet it is clearly known that various sources of uncertainty are present in all these phases. These uncertainties cannot be fully addressed solely by means of existing mechanistic knowledge or theory. Instead, a series of quality assurance exercises must be relied upon to reduce risks and to increase the foundation reliability. Reliability theory is a basis for formulating quality assurance (QA) plans and combining theoretic and observational information for decision-making. This paper outlines the important roles reliability theory and quality assurance play in each of these phases.

2 Characterization of uncertainties in foundation engineering

Geotechnical engineers face a great amount of uncertainty in foundation engineering. Some sources of uncertainty are inherent soil variability, loading effects, time effects,

construction effects, human errors, model errors, use of failure criteria, and random errors in soil boring, sampling, in-situ and laboratory testing, and characterization of strength and stiffness of soils. For a reliable design, the uncertainties must be identified, characterized, and taken into accounted in the design. Tang (1984, 1989), Wu et al. (1989), Christian et al. (1994), Kulhawy (1996), Duncan (2000), Zhang et al. (2001, 2004, 2005), Zhang and Ng (2005), Tang et al. (2006) and many others described principles for organizing and characterizing geotechnical uncertainties and presented excellent examples illustrating the use of probability theory to specifically include uncertainties in geotechnical design.

Tang (1984) characterized soil properties in a homogeneous soil layer for foundation performance evaluation. Three main sources of uncertainties including inherent spatial variability, statistical uncertainties due to limited number of soil samples and systematic uncertainties were identified. The systematic uncertainties included those contributed by the inability of a test to reproduce the in situ property due to factors such as sample disturbances, size of specimen, different stress conditions, etc and those in empirical calibration formula for estimating soil properties. A simple model was proposed to incorporate each of these uncertainties and relate the mean and coefficient of variation (COV) of the soil property of the tested soil specimens to those of the average in-situ soil property. Zhang et al. (2004) formalizes a procedure for characterizing and analyzing uncertainty in a design in which linear correlations have been used. The proposed procedure for reducing variability associated with an empirical correlation is summarized in Fig. 1. Having established a global empirical correlation based on a global dataset, the uncertainty associated with a design at a particular site using the correlation can be handled in three levels:

Level 1. No regional data and site-specific pile observation available. In this case, the uncertainty associated with the design using an empirical correlation is the uncertainty with the predicted value using the global empirical correlation. The prediction uncertainty or variance could be considerable because within-site variability, cross-site variability, and model errors are all involved. An example of the case is the correlations between the unit shaft resistance and toe resistance of piles and SPT blow count (Meyerhof 1976), which are often used directly in many parts of the world without considering specific local conditions.

Level 2. Regional data available. If a regional dataset is available and a regional correlation is derived, the predicted mean and variance based on the global correlation can be updated with those based on the regional correlation. The uncertainty in the prediction will then be reduced.

Level 3. Both regional data and site-specific observations available. In addition to reduction of variability using regional data as described in the second level, the uncertainty of prediction can be further reduced by including site-specific pile observations. Site-specific observation can lead to a significant reduction of variance. This provides a theoretic basis why proof tests and in-service inspections/monitoring should be an integral part of a construction process.

For the purpose of variance reduction using site-specific observations, Zhang et al. (2004) classified uncertainties into within-site variability and cross-site variability. Take pile load tests as an example. Suppose several "identical" test piles are constructed

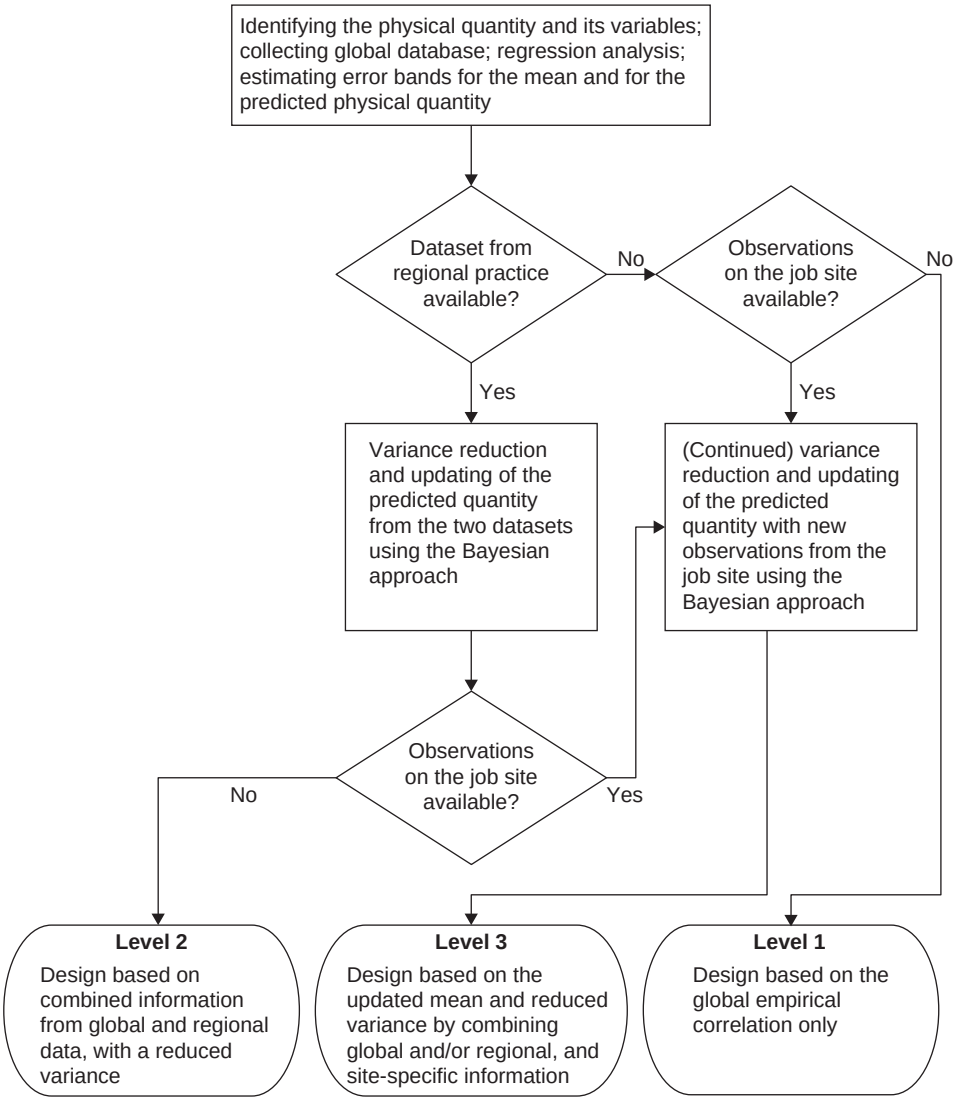


Figure 1 Proposed procedure for reducing variability with empirical correlations using regional and site-specific information (After Zhang et al. 2004).

at a seemingly uniform site and are load tested following an "identical" procedure. The measured values of ultimate bearing capacity of the piles would usually be different and the difference is due to the so-called "within-site" variability. The influence factors for the within-site variability of the pile capacity may include inherent variability of properties of the soil in the influence zone of each pile, construction effects, variability of pile geometry (length and diameter), variability of properties of the pile concrete,

and soil disturbance caused by pile driving and afterward set-up, and various construction effects. In general, design analyses are based on experiences from multiple sites. Cross-site variability is the additional variability arises from the differences in soils and pile materials, construction teams and construction details, quality control, and other factors at different sites. This variability can be greatly reduced based on site-specific observations.

3 Site investigation

In the site investigation phase, reliability theory is often used to characterize the spatial variability of geologic features. Geostatistics and random field theory (e.g. DeGroot and Baecher 1993, Fenten 1999) have been widely used to quantify the natural ground variability. More practically, reliability theory has been used to enhance the planning of investigation boreholes and in-situ tests so that any anomalies and key ground features can be revealed effectively. "Just one more boring, and we'll know for sure" (Liao et al. 1996) is a good example of the application of risk analysis to geotechnical exploration. Tang (1987) and Tang et al. (1989) analysed the probability of the presence of geotechnical anomalies and their probable size distribution. The effectiveness of using an extensive boring program in limiting the likelihood of the presence of an anomaly and, if present, in limiting the anomaly to acceptable size, was found to depend on boring grid patterns, the density of borings per unit area, and prior estimates of anomaly size and presence probabilities. The monetary value of any special site investigation plan can be evaluated in the reliability-based design framework based on the variability reduction from the investigation plan.

4 Design phase

In the design phase, evaluation of design analysis models and model parameters, as well as verification of any new construction workmanship are of great concern. A recommended way to risk management is to use an integrated approach that combines site investigation, design, and construction. In this approach, an early test program, for instance involving preliminary piles or early test piles, any required additional borehole boring, and a verification test program at the end of construction are implemented as means to manage various sources of uncertainty. These tests provide a means to verify design parameters and any new workmanship. The design can then be finalized according to outcomes of the early test program and verified by the verification tests at thee end of construction. A lower safety factor (FS) or higher resistance factor can be adopted, bearing in mind that the design will be verified at the end of construction. The benefit of such integrated approach can be quantified through a Bayesian framework of updating information.

4.1 *Example1-Direct application in design of barrettes*

A 480-m high tall building was constructed on a reclaimed ground. Due to the great depths of the bedrock (up to −130 m below the mean sea level), the building is supported on 240 barrettes with dimensions of 2.8 m × 1.5 m or 1.0 m. The major load baring layer is completely decomposed granite. Prior to the design of the foundation, there had been limited experience on the performance of barrettes in Hong Kong soils.

The following preliminary design parameters were recommended based on limited tests in previous projects (Ove Arup 2002):

(1) Ultimate shaft friction in alluvium = 6.25N, limited to 200 kPa, where N is the Standard penetration test (SPT) blow count;
(2) Ultimate shaft friction in CDG = 3.75N, limited to 200 kPa;
(3) Ultimate end-bearing in CDG = 10N, limited to 2000 kPa.

To verify the design soil parameters, five trial barrettes were constructed and load tested for determining shaft friction parameters in alluvium and the CDG soil layer for further detailed foundation design. These trial tests led to the following design parameters (Ove Arup 2002):

(1) Ultimate shaft friction in alluvium = 6.0N, limited to 200 kPa;
(2) Ultimate shaft friction in CDG = 2.4N, limited to 200 kPa;
(3) Ultimate end-bearing in CDG = 10N, limited to 2000 kPa.

The design of the production barrettes was finalized using the above parameters from the trial barrette tests. At the end of the construction of the production barrettes, four additional verification tests were conducted. Both the maximum settlements and the residual settlements of these verification tests were smaller than the respective allowable values; hence the performance of these barrettes was satisfactory.

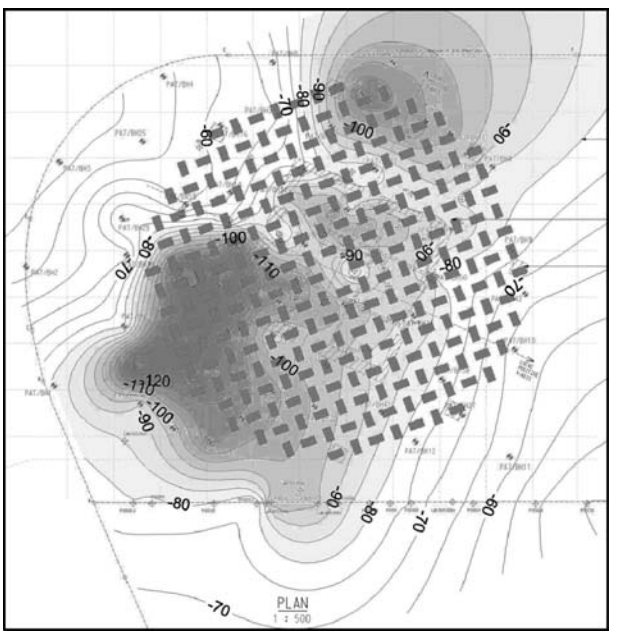


Figure 2 Design of barrettes based on trial and verification tests (After Ove Arup 2002).

4.2 *Example2-Design of driven piles*

In developing pile driving criteria for construction of driven piles, many types of construction control tests or analyses such as final set tests, dynamic tests using pile driving analyzers (PDAs), wave equation analyses with measurements, blow count and driving energy monitoring, and static loading tests are routinely carried out as a means to cope with uncertainties. These tests or analyses can be implemented in an integrated process. For instance, Bell et al. (2002) introduced a seven-step process for selecting driving criteria that has proven to be successful for the installation of driven piles. The process incorporates static pile capacity estimates, wave equation analyses, dynamic tests using PDA, and static loading tests in a rational and consistent sequence, and has resulted in reliable and cost-effective driving criteria. The steps of the sample process are:

(1) Evaluate subsurface data and perform static pile capacity analysis.
(2) Select preliminary driving criterion using wave equation analysis.
(3) Drive probe (or indicator) piles across the site area and use a PDA to evaluate capacity, driving stresses, and hammer performance, including re-tapping of selected piles after initial set-up.
(4) Evaluate and adjust the driving criterion based on the results of the PDA.
(5) Drive static test pile(s) using the revised driving criterion.
(6) Load test the test pile(s).
(7) Perform a final evaluation of the driving criterion for the production piles based on the static loading test results.

With these construction-control measures implemented, the risk level of the foundation being constructed should have been reduced. Accordingly, the factor of safety used in a conventional allowable stress design (ASD) can be reduced or the resistance factors in a reliability-based design (RBD) can be increased. Obviously, the degree the FS can be decreased depends on the level of construction control. This has long been recognized and practiced in the industry.

Table 1 summarizes the FS values commonly used by various design codes in conjunction with different levels of control during construction. When a reliable and consistent level of construction control is used, a smaller FS can be used. For example, the FS can be reduced from 3.50 for static analyses based on the standard penetration test (SPT) or the cone penetration test (CPT) to 2.00 for the same analyses that are verified with a sufficient number of static loading tests (ASSHTO, 1998). In AS-2159 (SAA, 1995), Eurocode 7 (CEN 2001) and Pile Driving Contractors Association (PDCA)'s specifications (PDCA 2001), the FS is even more specifically related to the number of construction control tests. Similarly in the reliability based design, the effect of construction control is also taken into account through modifying the values of resistance factor.

4.3 *An integrated approach*

The two aforementioned examples show how field observation during the early stage of construction can be utilized to assist cope with uncertainties. The effect of level of quality assurance can be quantified or the required safety factors or resistance factors to achieve a required reliability level can be derived rationally using reliability theory.

Table 1 Factor of safety and construction control (based on Likins 2004).

Construction control	*LRFDBDS (AASHTO 1998)*	*PDCA (2001)*	*Eurocode 7 (CEN 2001)*	*IBC (ICC 2000)*	*AS-2159 (SAA 1995)*
Static analysis	3.50	3.50	–	6.00	2.12–3.44
Dynamic formula	3.50	3.50	–	–	2.50–3.06
Wave equation	2.75	2.50	–	–	2.50–3.06
Dynamic tests with signal matching	2.25	–	–	2.00 [a]	–
Small number of tests	–	2.10 (2%)[b]	2.23 (2)[c]	–	2.12 (<3%)
Large number of tests	–	1.90 (10%)	1.95 (>20)	–	1.72 (>15%)
Static loading tests	2.00	–	–	2.00	–
Small number of tests	–	2.00 (<0.5%)	2.29 (1)	–	1.93 (<1%)
Large number of tests	–	1.65 (>5%)	1.64 (>5)	–	1.53 (>3%)
Static and dynamic tests [a]	1.90	–	–	–	–

Notes: (a) Requires at least one static loading test.
(b) Percentage of number of piles to be tested.
(c) Number of piles to be tested.

Zhang (2004) and Zhang et al. (2006b) proposed a procedure to integrate design and construction. The procedure provides a framework to investigate the effect of construction control on the safety of driven piles, and provides a theoretical basis to verify and support existing empirical factors of safety and resistance factors used to account for the effect of construction control. The procedure is formalized in the Bayesian framework. In particular, how the FS and resistance factor can be determined in a rational manner based on on-site dynamic pile tests is explored. The detailed formulations for the procedure can be found in Zhang et al. (2006b).

With necessary construction-control measures implemented, the risk level of the foundation being constructed should have been reduced. Accordingly, the factor of safety used in a conventional allowable stress design can be reduced or the partial factors or resistance factors in a limit state design can be reduced. Obviously, the degree the FS can be decreased depends on the level of construction control. This has long been recognized and practiced in the industry.

Take five methods for driven pile design as an example to illustrate the proposed procedure: two design methods for driven piles used in the design stage, the CPT method and Meyerhof's SPT method, and three methods used in the construction stage, wave equation analysis (WEAP) (GRL, 1995), Gates' formula (Gates, 1957), and Hiley's formula (Hiley 1922). The accuracy of a design method may be described by its mean bias factor and COV, and failure of a pile is defined by the Davisson criterion. The statistics of these design methods are presented in Table 2. The bias factors vary from 1.03 to 1.33 and the COV values vary from 0.35 to 0.50. Indeed, the ASD approach results in designs with levels of safety that are rather uneven from one method to another (i.e., $\beta = 1.92$–2.65).

Consider the case where a design is verified by one or several types of control tests, two tests for each type. The design is verified if the PDA capacity or the CAPWAP capacity is greater than the required capacity or the static loading tests pass designated acceptance criteria. A common level of target reliability, $\beta_T = 2.5$ for single piles is

Table 2 Bias factor and coefficient of variation for various design methods (After Zhang at al. 2006b).

Design method	*Number of cases*	*Bias factor*	*Coefficient of variation*	*ASD FS*	*Reliability index*	*References*
CPT method	–	1.03	0.36	2.50	1.98	Orchant et al. (1988)
SPT method	–	1.30	0.50	2.50	1.92	Meyerhof (1976)
WEAP	99	1.22	0.35	2.50	2.45	Rausche et al. (1996)
Gates' formula	122	1.33	0.48	3.00	2.40	Hannigan et al. (1997)
Hiley's formula	157	1.30	0.42	3.00	2.65	Zhang et al. (2006c)

Table 3 Required factor of safety after verification by different levels of construction control at $\beta_T = 2.5$ (After Zhang at al. 2006b).

Design method	*FS after verification by construction control*				
	PDA	*CAPWAP*	*Load test*	*PDA + Load test*	*CAPWAP + Load test*
CPT method	2.40	2.48	2.05	1.88	1.88
SPT method	2.24	2.15	2.03	1.80	1.80
WEAP	2.20	2.28	1.94	1.82	1.83
Gates' formula	2.18	2.11	1.98	1.81	1.77
Hiley's formula	2.18	2.17	1.97	1.80	1.77
Average	2.24	2.24	1.99	1.82	1.81

Note: (a) CAPWAP = PDA test with signal matching.
(b) The number of tests for all the construction control methods is two.

examined. Using the proposed procedure, the required FSs for designs after verification by various construction control methods or combinations of multiple construction control methods are obtained and summarized in Table 3.

The FS values recommended by ASSHTO (1998) when there is no any verification (Table 1) are used as a reference for comparison. It can be seen that

(1) The required FS can be substantially reduced through careful construction control. The required FS can be reduced from 3.50 for the static analysis method based on the SPT to 1.42 at $\beta_T = 2.0$, and to 1.80 at $\beta_T = 2.5$ for the same analysis that is verified with two CAPWAP analyses plus two static loading tests. which depends on design methods, on-site test methods, and the level of control implemented.
(2) The required FS decreases with the level of construction control. For instance, using the SPT method, the required FSs to reach $\beta_T = 2.5$ are 2.24, 2.03 and 1.80, respectively after verification by two PDA tests, two static loading tests, and two PDA tests plus two static loading tests.
(3) The required FSs for different design methods to reach the required target reliability are different because each method is associated with different bias and COV. Thus it is not reasonable to specify the same FS for different design methods if these design methods are not sufficiently verified by control testing.
(4) After verified by sufficient construction control testing, the required FSs for different design methods are not sensitive to the accuracy of the individual methods.

Table 4 Required resistance factor after verification by different levels of construction control at $\beta_T = 2.5$ (After Zhang at al. 2006b).

Design method	ϕ *after verification by construction control*				
	PDA	*CAPWAP*	*Load test*	*PDA + Load test*	*CAPWAP + Load test*
CPT method	0.57	0.55	0.66	0.72	0.72
SPT method	0.61	0.63	0.69	0.75	0.75
WEAP	0.62	0.60	0.70	0.75	0.75
Gates' formula	0.62	0.64	0.69	0.75	0.77
Hiley's formula	0.62	0.63	0.69	0.75	0.77
Average	0.61	0.61	0.69	0.74	0.75

For example, the required FSs for the five design methods in Table 2 after verification by two PDA tests plus two static loading tests fall into a narrow range of 1.80–1.88.

(5) The FSs obtained from this study are consistent with the existing empirical values as shown in Table 1. For $\beta_T = 2.5$, if the pile capacity is verified by two PDA tests or two CAPWAP analyses, the average FS calculated for the five design methods is 2.24. This is similar to that recommended in LRFDBDS (ASSHTO, 1998) and Eurocode 7 (CEN, 2001). The calculated average FS after verification by two static loading tests is 1.99, which is also similar to that in LRFDBDS (ASSHTO, 1998), and IBC (ICC, 2000). Similarly, if the pile capacity is first verified by two CAPWAP analyses and then by two static loading tests, the average required FS for the five design methods is 1.81, which is just slightly smaller than 1.90 in LRFDBDS (ASSHTO, 1998).

In the RBD, uncertainties with the pile capacity are considered using resistance factors. Construction control also influences the required resistance factor. For the five design methods studied, the required resistance factors to reach a target reliability index of 2.5 after verification by different construction control measures or combinations of multiple construction control measures are obtained and summarized in Table 4. It can be seen that

(1) The required ϕ increases with the level of construction control. For the design based on the SPT method, the required ϕ after verification by two PDA tests only, two static loading tests only, and two PDA tests plus two static loading tests are 0.61, 0.69, and 0.75, respectively.

(2) After satisfying sufficient construction control tests or analyses, the required ϕ values for the five design methods are not sensitive to the accuracy of the individual methods. For instance, after verification by two PDA tests plus two static loading tests, the required ϕ values for these design methods fall into a narrow range 0.72–0.75.

As in the LRFD Bridge Design Specifications (ASSHTO 1998), a factor λ_v can be used to take into account the effect of construction control in the RDB. In this study, the required resistance factor after verification by two CAPWAP analyses and then by two

Table 5 Calculated reduction factors for resistance factors at $\beta_T = 2.5$ after verification by different levels of construction control (After Zhang at al. 2006b).

Design method	*Construction control method*				
	PDA	*CAPWAP*	*Load test*	*PDA + Load test*	*CAPWAP + Load test*
CPT method	0.78	0.76	0.92	1.00	1.00
SPT method	0.80	0.84	0.89	1.00	1.00
WEAP	0.83	0.80	0.94	1.00	1.00
Gates' formula	0.81	0.84	0.89	0.98	1.00
Hiley's formula	0.81	0.82	0.90	0.98	1.00
Average	0.81	0.81	0.91	0.99	1.00

static loading tests is used as a reference to calculate λ_v. Table 5 shows the calculated values of λ_v at a reliability level of $\beta_T = 2.5$. These values are obtained by dividing the respective ϕ factors for each method by the ϕ factor that is sufficiently verified, for instance, by 0.72 for the CPT method. If the pile capacity is verified by two PDA tests and then by two static loading tests, the average λ_v for the five design methods is 0.99. The average λ_v after verification by two PDA tests or two CAPWAP analyses is 0.81. The effect of the number of control tests has been studied further by Zhang et al. (2006b).

5 Construction phase

The driven pile cases presented in the previous section show how quality assurances tests such as load tests and PDS tests, conducted during or at the end of construction, assist design and verification of foundations. Many quality assurance exercises such as construction supervision and routine quality control tests (e.g. various non-destructive integrity tests), however, do not provide direct evidence on pile performance. Rather these exercises, in the form of 'pass' or 'non-pass', are meant to check if a particular construction step or workmanship conforms to specified specifications. The impact of such routine quality assurance on pile reliability can be quantified using reliability theory. Planning of quality assurance schemes can also be rationally guided using reliability theory. These are illustrated using two examples with large-diameter bored piles.

5.1 *Application of probability theory to optimizing cross-hole integrity test specifications*

Large-diameter bored piles are commonly used to support heavy structures such as high-rise buildings and bridges. From time to time, due to errors in handling slurry, casings, reinforcement cages, concrete, and other factors, minor or major anomalies in piles can be introduced both during and after construction. An anomaly is any detected irregular feature. If an anomaly, because of either size or location, can affect pile shaft performance, then the anomaly is considered as a defect. Some examples of defects are voids, honeycombing, cracks, necking, soil inclusions, and corroded rebars (O'Neill and Reese 1999). The presence of these defects, especially if they are

sufficiently large and frequent, can lead to unsatisfactory performance of bored piles (Sarhan et al. 2004; O'Neill et al. 2003). Some quality assurance tests, such as cross-hole sonic logging (CSL), sonic echo, impedance logging, impulse response testing; interface or full-length coring, and static loading tests are routinely performed.

The reliability of CSL can be described by the inspection probability that an inspection finds a defect. The inspection probability is expressed as a product of the encountered probability and the detection probability (Li et al. 2005),

$$P(x) = P(E|x)P(D|E,x) \tag{1}$$

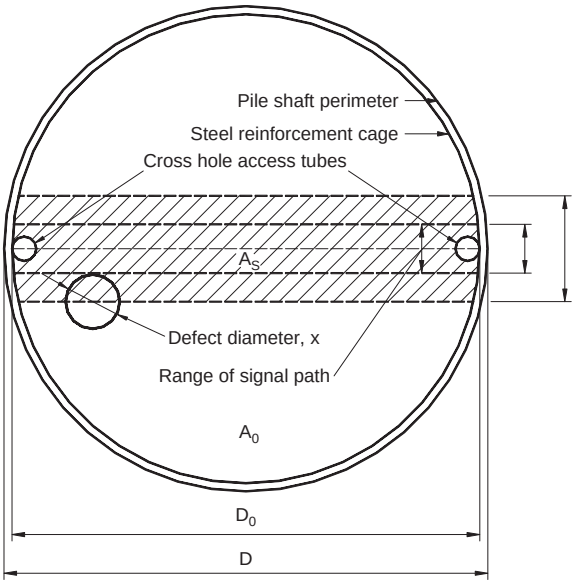
where $P(x)$ is the inspection probability for a given defect size x; E is the event that a defect with a given size x is encountered; D is the event that a defect with a given size x is detected if it is indeed encountered; $P(E|x)$ is the encountered probability that a defect is encountered by an inspection of a given inspection plan if a defect indeed exists; and $P(D|E, x)$ is the detection probability that an inspection detects a defect if a defect is indeed encountered. To conduct the reliability evaluation of CSL, it is necessary to determine $P(E|x)$ and $P(D|E, x)$.

Defects present in a pile may not be encountered during a CSL test or a interface coring. Even when defects are indeed encountered, they may not always be detected. Consider first the case where a defect with a diameter of x is present in a pile with two access tubes as shown in Fig. 3. Theoretically, a cross-hole sonic logging test can scan the path between the two access tubes with a width of about 2λ, where λ is the wavelength of the ultrasonic signal (Hassan and O'Neill 1998). The smallest defect size that can be reasonably detected by CSL is about one wavelength of the ultrasonic signal (Finno and Champy 1997). The wave velocity, v, in good quality concrete is typically about 4000 m/s (O'Neill and Reese 1999), so the wavelength is $\lambda = v/f = 0.08$ m. Under the above assumptions, if the signal path encounters a defect, the center of the defect must be located inside the shaded area. In other words, the distance between the center of the defect and the signal path must be larger than the radius of the defect. The probability of the event that a defect can be encountered by any of the signal paths can be simply taken as the ratio of the shaded area to the cross-sectional area of reinforcement cage,

$$P_E(E|x) = \frac{A_S}{A_0} \tag{2}$$

where A_0 and A_S are the shaded area and the cross-sectional area of reinforcement cage, respectively; A_0 can be calculated in terms of the diameter of the reinforcement cage D_0, which is approximately equal to the pile diameter D for large-diameter bored piles.

Table 6 shows some recommended numbers of access tubes for given defect sizes as a percentage of pile cross-sectional area, p, at encounter probability values of $P_E = 0.9$ and $P_E = 0.95$. The results justify some common practices shown in Table 7 and help avoid the use of excessive access tubes when the pile diameter is large. When results from on-site coring are available, the estimated occurrence probability and defect size can be updated in the Bayesian framework. The planning of the types and number of tests in a QA program can be assisted by a cost-benefit analysis (Li et al. 2005).



A_0 = Cross-sectional area of reinforcement cage.
A_s = Pile area over which CSL test operates.
D = Pile diameter.
D_0 = Diameter of reinforcement cage.
λ = Wave length.

Figure 3 Configuration of two access tubes over the cross section of a pile.

Table 6 Recommended number of access tubes for given defect size as a percentage of pile cross-sectional area *p* at encounter probability values of $P_E = 0.9$ and $P_E = 0.95$ (After Li et al. 2005).

Pile diameter (mm)	*p = 5%*		*p = 10%*		*p = 15%*	
	$P_E = 0.9$	$P_E = 0.95$	$P_E = 0.9$	$P_E = 0.95$	$P_E = 0.9$	$P_E = 0.95$
600–750	3	3–4	3	3	3	3
750–1000	3–4	4	3	3	3	3
1000–1500	4	4	3	3–4	3	3
1500–2000	4	4	3	4	3	3
2000–2500	4	4	3–4	4	3	3
2500–3000	4	4	4	4	3	3

A general expression for the detection probability for a given size x is obtained as

$$P_D(x) = \begin{cases} 0 & 0 < x \leq 200 \text{ mm} \\ \dfrac{x - 200}{100} & 200 < x \leq 300 \text{ mm} \\ 1 & x > 300 \text{ mm} \end{cases} \quad (3)$$

Table 7 Summary of recommended number of access tubes for different pile diameters.

Pile diameter (mm)	*Tijou (1984)*	*Turner (1997)*	*O'Neill and Reese (1999)*	*Works Bureau (2000)*	*Thasnanipan et al. (2000)*	*JGJ 106–2003 (MOC 2003)*
600–750	2	3	2	3–4	2	2
750–1000	2–3	3–4	2–3	3–4	3	2–3
1000–1500	4	4–5	4–5	3–4	4	3
1500–2000	4	4–5	5–7	3–4	6	3
2000–2500	4	4–5	7–8	3–4	6	4
2500–3000	4	4–5	8	3–4	8	4

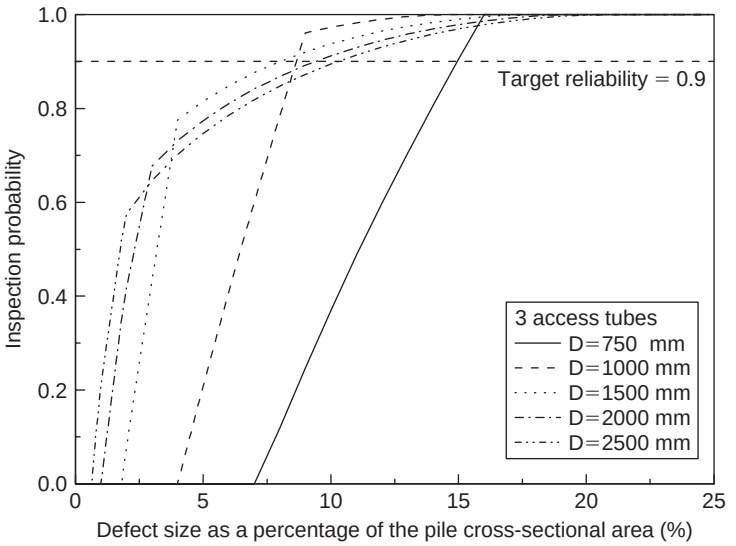


Figure 4 Inspection probability for defect size as a percentage of the pile cross-sectional area (After Li et al. 2005).

where the minimum detectable crack size below which a crack cannot be detected is taking as 200 mm and the detectable crack size with certainty beyond which a crack can be detected with certainty is taken as 300 mm. Combining eqs. (2) and (3), the inspection probability for anomalies smaller than a certain size can be determined. Fig. 4 illustrates the inspection probability for defect size as a percentage of the pile cross section area.

5.2 *Reliability and impact of routine quality assurance for bored piles*

Quality assurance tests, such as integrity tests, are routinely conducted to ensure the safety of pile foundations. QA tests provide additional information and result in changes in estimated reliability of pile foundations. Zhang et al. (2006a) formalized a procedure to quantitatively evaluate the impact of routine QA tests on the reliability of pile foundations. The prior information of the occurrence probability and thickness

of toe debris is established based on a practice survey and accumulated QA data. The Bayesian approach is then applied to update the occurrence probability and the mean toe debris thickness based on outcomes of on-site QA tests and remedial actions taken after these QA tests. Subsequently, the reliability of the piles can be updated.

Consider the case of single bored piles with toe debris along with other types of imperfections such as cracks, necking, and voids. If thick debris exists, then the pile capacity or settlement will be adversely affected. The likelihood of toe debris presence in a pile depends on construction conditions, construction procedure, results of interface coring, and engineers' judgments. The toe debris, if present, is assumed to be uniformly distributed over the pile cross section for simplicity. The pile performance is still uncertain even when the pile is free from toe debris due to the presence of other types of imperfections as well as many other sources of uncertainty such as spatial variability of soils, various construction effects, and load effects. Based on the total probability theorem (e.g., Ang and Tang 2007), the probability of unsatisfactory performance of the pile, p_f, can be calculated as follows:

$$p_f = P(F|\overline{E})(1 - p_d) + P(F|E)\, p_d \tag{4}$$

where p_d = occurrence probability of toe debris; F = event of pile failure; E and $\overline{E}$ = events of toe debris presence and toe debris absence, respectively; $P(F|\overline{E})$ and $P(F|E)$ = conditional probabilities of unsatisfactory performance of the pile given the absence and presence of toe debris, respectively.

For a particular pile, the toe debris may be described by its thickness x. If toe debris does exist and x is taken to be a continuous variable, its probability distribution can be expressed as $f(x|t)$ in which the parameter, t, is the mean of x and treated as another random variable with a probability distribution $f(t)$. The conditional probability of unsatisfactory performance of the pile can be further given by

$$P(F|E) = \int_{x_L}^{x_U} P(F|x)\left[\int_{t_L}^{t_U} f(x|t) f(t)\, dt\right] dx \tag{5}$$

where x_L and x_U = lower and upper bounds of x; t_L and t_U = lower and upper bounds of t.

For a pile that is constructed without any QA testing, the empirical occurrence probability of toe debris can be considered as a prior distribution of p_d. The empirical perception of toe debris thickness can also be used to establish the prior distribution of toe debris thickness. A prior p_f can then be calculated using Eq. (4). If field measurements from interface coring are available, the distributions of p_d and toe debris thickness can be updated with the additional measurement information using the Bayesian approach (e.g., Ang and Tang 2007), and $P(F|E)$ of the pile with toe debris can be calculated using Eq. (5). Once the updated distribution of p_d and the updated conditional reliability of the pile with toe debris are available, the updated p_f can be calculated using Eq. (4). The following three scenarios are considered in this study: no toe debris detected, toe debris detected without repair, and toe debris detected and repaired.

5.2.1 *Case 1 – No toe debris detected*

Suppose n piles are selected randomly from a site for interface coring. If the construction quality at the site is excellent and no toe debris is found in the sampled piles, then no actions need to be taken after the QA tests. In this case, the occurrence probability of toe debris will be updated after the QA tests, and the updated occurrence probability will be smaller than the estimated prior value. Since no information on the thickness of toe debris is obtained, the distribution of thickness of toe debris cannot be updated. According to Eq. (4), the updated probability of unsatisfactory performance of the pile will be smaller than that before the QA tests. The reliability of the piles will be improved though remedial actions are not taken.

5.2.2 *Case 2 – Toe debris detected without repair*

Suppose n piles are selected randomly from a site for inspection using the interface coring and m out of these are found to contain toe debris. Then, a decision on immediate repair or replacement must be made. Let's assume the detected amount of toe debris is tolerable so that no repair actions are necessary. In this case, on-site information on both the occurrence probability and the toe debris thickness can be obtained from the QA tests. Accordingly, the distributions of occurrence probability and toe debris thickness can both be updated after the QA tests. Thus, the p_f of the pile can be updated. In particular, if the observed occurrence probability is smaller than the prior occurrence probability, and the observed toe debris thickness is less than the prior thickness, then the updated p_f after the QA tests will be smaller than that before the QA tests. The reliability of the piles will then be changed even though remedial actions such as repair are not taken after toe debris is detected.

5.2.3 *Case 3 – Toe debris detected and repaired*

Suppose some toe debris is found in a random sampling at a site as in case 2. If the detected toe debris is deemed intolerable, the toe debris needs to be repaired (e.g., by pressure grouting) or the defective piles need to be replaced. Upon detecting serious toe debris in the sampled piles, further inspection on other piles at this site may be carried out. If it is further assumed that all toe debris present will be detected and repaired, the occurrence probability of toe debris becomes zero. Obviously, the updated p_f will be significantly smaller than that before the QA tests and the repair have been conducted.

Table 8 summarizes the parameters that can be and cannot be updated through the interface coring tests.

McVay et al. (1998) calibrated the design and construction procedures outlined by Reese and O'Neill (1988) and McVay et al (1992) for bored piles founded on rock. When the soil shaft resistance, the rock shaft resistance, and the end bearing are all considered, the bias factor and COV of the pile capacity are $\lambda_R = 1.15$ and $COV_R = 0.17$. The reliability of such piles can thus be calculated. The test piles in the calibration exercise are assumed to have been properly constructed. If excessive toe debris exists, then λ_R will decrease due to the reduction in the pile capacity.

The effect of toe debris of varying thickness on the behavior of large-diameter bored piles with different lengths can be analyzed using a nonlinear finite-element method. For bored piles of three pile lengths (25, 50, and 75 m) supported in weathered granitic

Table 8 Summary of updated items for all three cases.

Case	*Assumptions*	*PDF of occurrence probability of toe debris, $f(p_d)$*	*PDF of toe debris thickness, $f(t)$*
Case 1	No toe debris detected	Updated	Not updated
Case 2	Toe debris detected but not repaired	Updated	Updated
Case 3	Toe debris detected and repaired	$p_d = 0, x = 0$	Not relevant

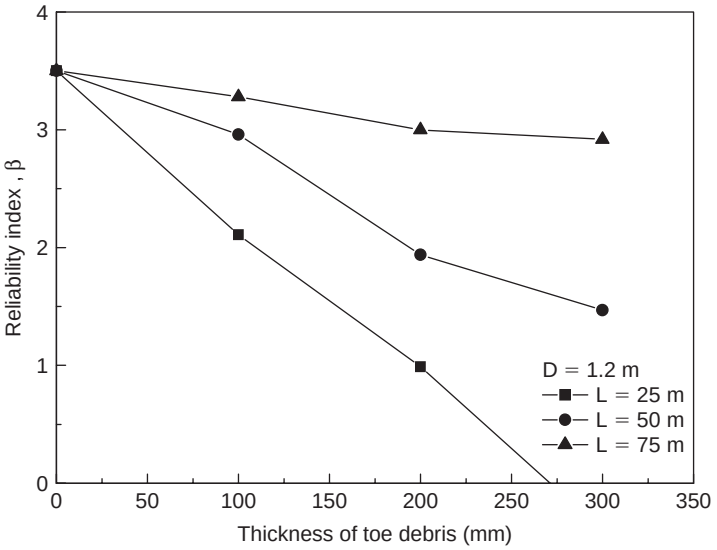


Figure 5 Reliability of piles given toe debris of varying thickness (After Zhang et al. 2006a).

soils underlain by granitic rocks, Zhang et al. (2006a) obtained the relationships between the reliability index and the toe debris thickness, as shown in Fig. 5. The reliability index decreases significantly due to the effect of toe debris when the bedrock levels are not extremely deep (say, smaller than 50 m). For instance, for a 1.2 m diameter, 25 m long pile, the reliability index decreases from 3.5 when no toe debris is present to 2.11 when a 100-mm thick toe debris is present. The toe debris has a minor effect on the 75 m long piles. Obviously, when a pile is very long, the mobilized toe resistance only consists of a small fraction of the pile capacity. The reliability index at zero toe debris thickness, 3.50, and the corresponding probability of failure, $P(F|\overline{E}) = 2.33 \times 10^{-4}$, account for other types of imperfections besides toe debris as well as many other sources of uncertainty.

The conditional probability of unsatisfactory performance given the presence of toe debris, $P(F|E)$, can be calculated using Eq. (5). To do so, three functions, $P(F|x)$, $f(x|t)$, and $f(t)$ must be available. $P(F|x)$ can be obtained by curve fitting based on the data

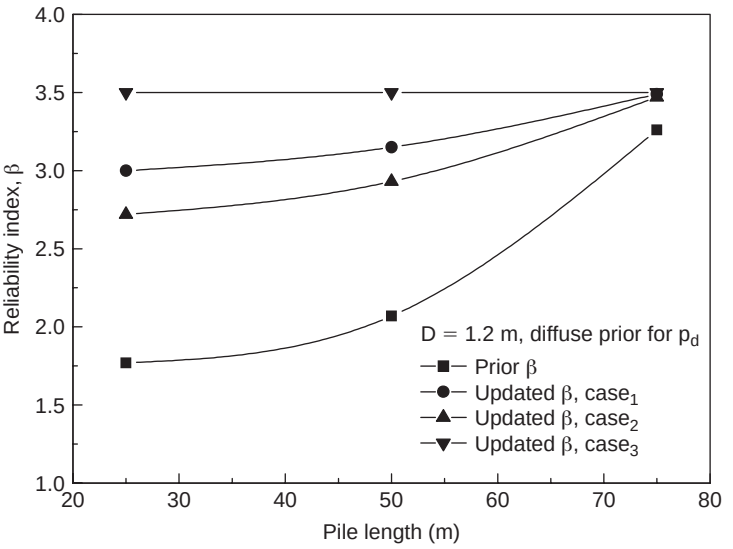


Figure 6 Comparison between updated and prior reliability indexes of piles using a diffuse prior for p_d (After Zhang et al. 2006a).

in Fig. 5; $f(x|t)$ is expressed by an exponential distribution; and $f(t)$ is expressed by the inverted gamma distribution.

Having determined $P(F|E)$, $P(F|\overline{E})$, and the Bayesian estimators of p_d, the reliability of the piles at a site before and after the QA tests can be calculated by Eq. (4) for each of the three cases of QA test outcomes and follow-up actions described earlier. Fig. 5 shows the reliability of the 1.2 m diameter piles of different lengths before and after QA tests for the scenario of a diffuse prior p_d. If no toe debris is detected in the QA tests, the distribution of the occurrence probability is significantly modified and the uncertainty in the occurrence probability is substantially reduced. Accordingly, the reliability of the pile will be increased. If toe debris is detected but no repair actions are taken, the distributions of both the occurrence probability and the thickness of toe debris will still be updated. The estimated reliability of the piles will then be changed. If all toe debris is detected and repaired, the occurrence probability will become zero, and the reliability of the piles may be increased substantially.

The QA test information can improve the reliability of pile foundations significantly in the example. If no prior information on occurrence probability is available, the reliability index of the 1.2 m diameter, 25 m long piles can be increased from the prior value of 1.77 to 3.00, 2.72, and 3.50, respectively in the 3 cases. The degree of reliability improvement depends on outcomes of the QA tests, actions taken after QA tests, the pile length, and the pile diameter. Prior information, or engineers' experience, can play an important role in handling uncertainties in pile construction. However, the prior information may be associated with large uncertainties and overruled by site-specific QA test information.

6 Summary

Quality assurance and reliability theory play an important role in managing construction risks in all three phases of foundation design and construction: site investigation, design, and construction. The role may be summarized as follows:

(1) Site investigation is the most important step to reveal and manage construction risks. In the site investigation phase, reliability theory is used to characterize the spatial variability of geologic features, to optimize the planning of investigation boreholes and in-situ tests so that any anomalies and key ground features can be revealed more effectively, and to evaluate the cost effectiveness of any special site investigation plan.
(2) In the design phase, an recommended approach to risk management is to integrate site investigation, design, and construction. In this approach, an early test program, for instance involving preliminary piles or early test piles, necessary additional borehole boring, and a verification test program at the end of construction are implemented as means to manage various sources of uncertainty (e.g. verifying and finalizing design models, design parameters, and any new construction workmanship). The integration can be formulated within a Bayesian framework of updating information. With this integrated approach, costs saving may be achieved while maintaining the risks within an acceptable level.
(3) In the construction phase, effective risk management and decision-making can be made possible by combining both quantitative and qualitative information from quality assurance exercises using reliability theory. In particular, it is valuable to apply the Bayesian method to quantify the qualitative information (e.g. from site supervision and integrity tests in the form of "pass" or "no anomaly" statements) for updating design parameters and assisting decision-making.

Acknowledgements

The research described in this paper is substantially supported by the Research Grants Council of the Hong Kong SAR (Projects Nos. HKUST6035/02E and HKUST6126/03E).

References

American Association of State Highway and Transportation Officials (AASHTO), *LRFD Bridge Design Specifications*, 2nd edition, Washington, D.C., 1998.

Ang, A.H.–S. & Tang, W.H., *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, Vol. 1, 2nd edition, Wiley, New York, 2007.

Bell, K.R., Davie, J.R., Clemente, J.L. & Likins, G., Proven Success for Driven Pile Foundations, *Proc. Int. Deep Foundations Congress, Geotech. Spec. Publ. No. 116* (ed. by O'Neill, M.W. and Townsend, F.C.), ASCE, Reston, Va., pp. 1029–1037, 2002.

CEN, *Eurocode 7-Geotechnical Design, Part 1 General Rules*, European Committee for Standardization (CEN), 2001.

Christian, J.T., Ladd, C.C. & Baecher, G.B., Reliability Applied to Slope Stability Analysis, *J. Geotech. Engrg.,* ASCE, Vol. 120, No. 12, pp. 2180–2207, 1994.

DeGroot, D.J. & Baecher, G.B., Estimating Autocovariance of In-situ Soil Properties, *Journal of Geotechnical Engineering,* Vol. 119, No. 1, pp. 147–166, 1993.

Duncan, J.M., Factors of Safety and Reliability in Geotechnical Engineering, *J. Geotech. and Geoenvir. Engrg.*, ASCE, Vol. 126, No. 4, pp. 307–316, 2000.

Fenton, G.A., Estimating for Stochastic Soil Models, *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 125, No. 6, pp. 470–485, 1999.

Finno, R.J. & Gassman, S.L., Impulse Response Testing of Drilled Shafts, *J. Geotech. Geoenviron. Eng.,* Vol. 124, No. 10, pp. 965–975, 1998.

Gates, M., Empirical Formula for Predicting Pile Bearing Capacity, *Civil Engineering*, Vol. 27, No. 3, pp. 65–66, 1957.

GRL and Associates, Inc., *GRLWEAP Manual*, 4535 Emery Industrial Parkway, Cleveland, Ohio 44128, USA, 1995.

Hannigan, P.J., Goble, G.G., Thendean, G., Likins, G.E. & Rausche, F., *Design and Construction of Driven Pile Foundations,* Workshop Manual, Vol. 1, *Publication No. FHWA-HI-97-014*, Federal Highway Administration, Washington, D.C., 1997.

Hassan, K.M. & O'Neill, M.W., *Structural Resistance Factors for Drilled Shafts with Minor Defects*, Final Rep., Phase I, Dept. of Civil and Environmental Engineering, Univ. of Houston, Houston, 1998.

Hiley, A., The Efficiency of the Hammer Blow and Its Effects with Reference to Piling, *Engineering*, Vol. 2, pp. 673, 1922.

ICC, *International Building Code*, International Code Council (ICC), USA, 2000.

Kulhawy, F.H., From Casagrande's "Calculated Risk" to Reliability-based Design in Foundation Engineering, *J. Boston Society of Civil Engineers*, Vol. 11, No. 2, pp. 43–56, 1996.

Li, D.Q., Zhang, L.M. & Tang, W.H., Evaluation of Reliability of Cross-hole Sonic Logging for Bored Piles, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 131, No. 9, pp. 1130–1138, 2005.

Liao, S.S.C., Druss, D.L., Neff, T.L. & Brenner, B.R., Just One More Borehole, and We'll Know for Sure. *Uncertainty in the Geologic Environment: From Theory to Practice.* Geotechnical Special Publication No. 58, Vol. 2, ASCE, New York, pp. 119–133, 1996.

Likins, G.E., Pile Testing-selection and Economy of Safety Factors, *Current Practices and Future Trends in Deep Foundations, Geotech. Spec. Publ. No. 125* (ed. by DiMaggio, J. A. and Hussein, M. H.), ASCE, Reston, Va., pp. 239–252, 2004.

Mcvay, M.C., Kuo, C.L. & Singletary, W.A., *Calibrating Resistance Factors in the Load and Resistance Factor Design for Florida Foundations*, Final Rep., Dept. of Civil Engineering, Univ. of Florida, Gainesville, 1998.

McVay, M.C., Townsend, F.C. & Williams, R.C., Design of Socketed Drilled Shafts in Limestone, *J. Geotech. Eng.*, Vol. 118, No. 10, pp. 1626–1637, 1992.

Meyerhof, G.G., Bearing Capacity and Settlement of Pile Foundations, *J. Geotech. Engrg.*, Vol. 102, No. 3, pp. 195–228, 1976.

Ministry of Construction (MOC), *Technical Code for Testing of Building Foundation Piles (JGJ 106–2003),* Beijing, China, 2003.

O'Neill, M.W. & Reese, L.C., *Drilled Shafts: Construction Procedures and Design Methods*, *Publ. No. FHWA-IF-99-025*, Federal Highway Admin., Office of Implementation, McLean, Va, 1999.

O'Neill, M.W., Tabsh, S.W. & Sarhan, H. A., Response of Drilled Shafts with Minor Flaws to Axial and Lateral Loads, *Eng. Struc.*, Vol. 25, No. 1, pp. 47–56, 2003.

Orchant, C.J., Kulhawy, F.H. & Trautmann, C.H., Reliability-based Foundation Design for Transmission Line Structures, Vol. 2: Critical Evaluation of In Situ Test Methods, *EL-5507 Final report*, Electrical Power Institute, Palo Alto, Calif., 1988.

Ove Arup and Partners Hong Kong Ltd., Construction of Barrettes and Diaphragm Walls at Union Square, Hong Kong, 2002.

PDCA, *Recommended Design Specifications for Driven Bearing Piles*, Pile Driving Contractors Association (PDCA), Boulder, CO, USA, 2001.

Rausche, F., Thendean, G., Abou-Matar, H., Likins, G.E. & Goble, G.G., Determination of Pile Driveability and Capacity from Penetration Tests, *FHWA Report DTFH61-91-C-00047*, Federal Highway Administration, Washington, D. C. , 1996.

SAA, Australian Standards: *Piling – Design and Installation AS-2159*, Standards Association of Australia (SAA), Homebush, NSW, Australia, 1995.

Sarhan, H.A., O'Neill, M.W. & Tabsh, S.W., Structural Capacity Reduction for Drilled Shafts with Minor Flaws, *ACI Struct. J.,* Vol. 101, No. 3, pp. 291–297, 2004.

Tang, W.H., Principles of Probabilistic Characterization of Soil Properties, *Probabilistic Characterization of Soil Properties: Bridge between Theory and Practice*, D.S. David and H.–Y. Ko (Eds.), ASCE, New York, pp. 74–89, 1984.

Tang, W.H., Updating Anomaly Statistics – Single Anomaly Case, *Structural Safety*, No. 4, pp. 151–163, 1987.

Tang, W.H., Uncertainties in Offshore Axial Pile Capacity, *Geotechnical Special Publication*, No. 27, Vol. 2, ASCE, New York, pp. 833–847, 1989.

Tang, W.H. & Gilbert, R.B., Case Study of Offshore Pile System Reliability. *Proc., 25th Offshore Technology Conference*, Houston, TX, pp. 677–686, 1993.

Tang, W.H., Sidi, I. & Gilbert, R.B., Average Property in a Random Two-state Medium, *Journal of Engineering Mechanics*, ASCE, Vol. 115, No. 1, pp. 131–144, 1989.

Tang, W.H, Zhang, L.M. & Zheng, Y.R., Dealing with Uncertainty in Engineering Design for Large-scale Gravel Soil Slopes in the Three Gorges Reservoir Zone, Keynote lecture paper, *Geohazards – Technical, Economical and Social Risk Evaluation,* 18–21 June 2006, Lillehammer, Norway. Farrokh Nadim (ed.), Berkeley Electronic Press, in CD Rom.

Thasnanipan, N., Maung, A.W., Navaneethan, T. & Aye, Z.Z., Non-Destructive Integrity Testing on Piles Founded in Bangkok Subsoil, *Proc., 6th Int. Conf. on the Application of Stress-wave Theory to Piles*, S. Niyama and J. Beim, eds., A.A. Balkema, Rotterdam, pp. 171–177, 2000.

Tijou, J.C., Integrity and Dynamic Testing of Deep Foundations – Recent Experiences in Hong Kong (1981–83), *Hong Kong Engineer*, Vol. 12, No. 9, pp. 15–22, 1984.

Turner, M.J., *Integrity Testing in Piling Practice*, Construction Industry Research and Information Association, CIRIA Report 144, London, 1997.

Works Bureau, *Enhanced Quality Supervision and Assurance for Pile Foundation Works,* Works Bureau Technical Circular No. 22/2000, Hong Kong, 2000.

Wu, T.H., Tang, W.H., Sangrey, D.A. & Baecher, G.B., Reliability of Offshore Foundations-State of the Art, *J. Geotech. Engrg. Div*., ASCE, Vol. 115, No. 2, pp. 157–178, 1989.

Zhang, L.M., Reliability Verification Using Proof Pile Load Tests, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 130, No. 11, pp. 1203–1213, 2004.

Zhang, L.M. & Ng, A.M.Y., 2005, Probabilistic Limiting Tolerable Displacements for Serviceability Limit State Design of Foundations, *Geotechnique*, Vol. 55, No. 2, pp. 151–161, 2005.

Zhang, L.M., Li, D.Q & Tang, W.H., Reliability of Bored Pile Foundations Considering Bias in Failure Criteria, *Canadian Geotech. Journal*, Vol. 42, No. 4, pp. 1086–1093, 2005.

Zhang, L.M., Li, D.Q. & Tang, W.H., Impact of Routine Quality Assurance on Reliability of Bored Piles, *J. Geotech. Geoenviron. Engrg*., ASCE, Vol. 132, No. 5, pp. 622–630, 2006a.

Zhang, L.M., Li, D.Q. & Tang, W.H., Level of Construction Control and Safety of Driven Piles, *Soils and Foundations*, Vol. 46, No. 4, pp. 415–426, 2006b.

Zhang, L.M., Shek, M.P., Pang, W.H. & Pang, C.F., Knowledge-based Pile Design using a Comprehensive Database. *Geotech. Engrg., Proc. of the Institution of Civil Engineers*, Vol. 159, No. GE3, pp. 177–185, 2006c.

Zhang, L.M., Tang, W.H. & Ng, C.W.W., Reliability of Axially Loaded Driven Pile Groups, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 127, No. 12, pp. 1051–1060, 2001.

Zhang, L.M., Tang, W.H., Zhang, L.L. & Zheng, J.G., Reducing Uncertainty of Prediction from Empirical Correlations, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 130, No. 5, pp. 526–534, 2004.

Chapter 6

Geosynthetics for soil reinforcement

R.D. Holtz

Department of Civil & Environmental Engineering, University of Washington, Seattle, WA, USA

ABSTRACT: The paper introduces geosynthetic materials and briefly describes their types and manufacture, functions and applications, properties and tests, design, selection, and specifications. Then the three primary soil reinforcement applications using geosynthetics – embankments on very soft foundations, increasing the stability of steep slopes, and reducing the earth pressures behind retaining walls and abutments – are discussed in some detail. Emphasis is on the properties of the geosynthetics required for design and construction.

1 Introduction

Historically, major developments in structural engineering have only been possible because of parallel developments in the technology of construction materials. Larger and more elaborate structures became possible as we progressed from using wood to building stone to concrete to reinforced concrete and most recently to prestressed reinforced concrete.

The development of steel enabled the construction of longer span bridges and taller buildings than were possible using, for example, wrought iron. Because the materials of geotechnical engineering are soil and rock, it is difficult to think of similar parallel developments in geotechnical construction and earthen materials in our field. Compaction and other soil improvement techniques occurred largely because of innovations in construction equipment by manufacturers and contractors. Probably the best example of a parallel development between a material and the construction application is soil reinforcement. In a direct analogy with reinforced concrete, steel and polymeric materials provide tensile resistance and stability to soils that have low to no tensile strength.

Polymeric reinforcement materials are a subset of a much larger recent development in civil engineering materials: *geosynthetics*. Geosynthetics are planar products

* Updated from Holtz, R.D. (2001) "Geosynthetics for Soil Reinforcement", *Materials Science for the 21st Century*, Proceedings of the 50th Anniversary Meeting, The Society of Materials Science Japan, Invited Papers, Volume A, pp. 162–171. Also given as the Ninth Spencer J. Buchanan Lecture, Texas A&M University, 9 November 2001, pp. 12–30 and as a paper in Current Trends in Geotechnical Engineering, Proceedings of the 19th Central Pennsylvania Geotechnical Conference (2002), ASCE/PennDoT, Hershey, Pennsylvania, Paper 6, pp. 1–20. A version of the paper was also presented at the XXXIV Ohio River Valley Soils Seminar, Lexington, Kentucky, 19 September 2003, as a keynote lecture at the International Conference on Geotechnical Engineering, University of Shajah, UAE, October 6, 2004, and at the 23rd Annual Geotechnical Seminar GEO-Omaha 2006, February 17, 2006, Omaha, Nebraska.

manufactured from polymeric materials (the *synthetic*) used with soil, rock, or other geotechnical-related material (the *geo*) as part of a civil engineering project or system. There are few developments that have had such a rapid growth and strong influence on so many aspects of civil engineering practice as geosynthetics. In 1970, there were only five or six geosynthetics available, while today more than 600 different geosynthetic products are sold throughout the world. The size of the market, both in terms of square metres produced and their value, is indicative of their influence. Worldwide annual consumption of geosynthetics is about 1.5×10^6 m^2, and the value of these materials is probably US\$$3 \times 10^6$. Since the total cost of the construction is at least four or five times the cost of the geosynthetic itself, the impact of these materials on civil engineering construction is very large indeed.

In less than 30 yr, geosynthetics have revolutionized many aspects of our practice, and in some applications they have entirely replaced traditional construction materials. In many cases, a geosynthetic can significantly increase the safety factor, improve performance, and reduce costs in comparison with conventional design and construction alternates.

This paper begins with a brief introduction to geosynthetic materials, with a description of the types and manufacture, functions and applications, properties and tests, design, selection, and specifications of geosynthetics. The remainder of the paper deals with the use of geosynthetics for soils reinforcement, with specific applications to embankments on soft foundations, steep slopes, and retaining walls and abutments.

2 Definitions, types, manufacture, and identification

2.1 *Definitions and types*

ASTM has defined a *geosynthetic* as a planar product manufactured from a polymeric material used with soil, rock, earth, or other geotechnical-related material as an integral part of a civil engineering project, structure, or system. A *geotextile* is a permeable geosynthetic made of textile materials. *Geogrids* are primarily used for reinforcement; they are formed by a regular network of tensile elements with apertures of sufficient size to interlock with surrounding fill material. *Geomembranes* are low-permeability geosynthetics used as fluid barriers. Geotextiles and related products such as nets and grids can be combined with geomembranes and other synthetics to take advantage of the best attributes of each component. These products are called *geocomposites*, and they may be composites of geotextile-geonets, geotextile-geogrids, geotextile-geomembranes, geomembrane-geonets, geotextile-polymeric cores, and even three-dimensional polymeric cell structures. There is almost no limit to the variety of geocomposites that are possible and useful. The general generic term encompassing all these materials is *geosynthetic*. A convenient classification system for geosynthetics is given in Fig. 1.

2.2 *Types and manufacture*

Most geosynthetics are made from synthetic polymers such as polypropylene, polyester, polyethylene, polyamide, PVC, etc. These materials are highly resistant to biological and chemical degradation. Natural fibers such as cotton, jute, bamboo, etc., could be used as geotextiles and geogrids, especially for temporary applications, but with

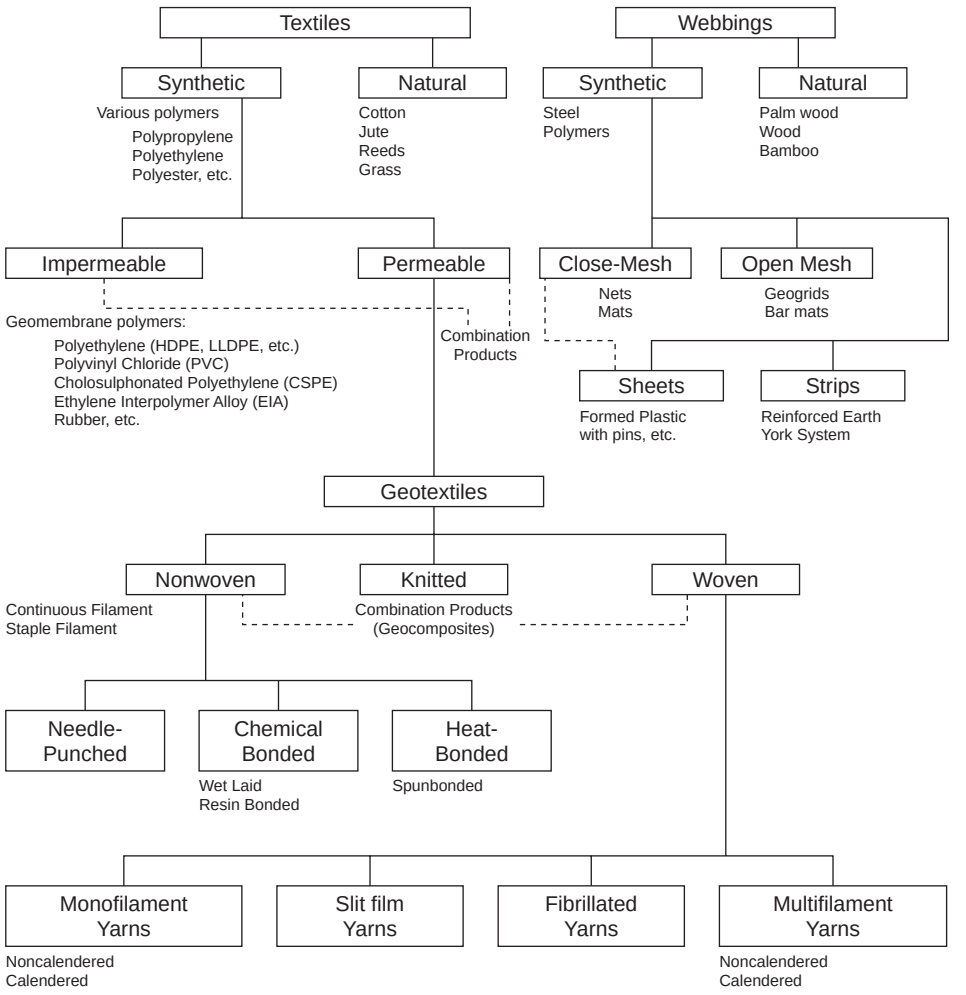


Figure 1 Classification of geosynthetics and other soil inclusions.

few exceptions they have not been promoted or researched as widely as polymeric geosynthetics.

In manufacturing geotextiles, elements such as fibers or yarns are combined into planar textile structures. The fibers can be continuous filaments, which are very long thin strands of a polymer, or staple fibers, which are short filaments, typically 20 to 100 mm long. The fibers may also be produced by slitting an extruded plastic sheet or film to form thin flat tapes. In both filaments and slit films, the extrusion or drawing process elongates the polymers in the direction of the draw and increases the fiber strength.

Geotextile type is determined by the method used to combine the filaments or tapes into the planar textile structure. The vast majority of geotextiles are either *woven* or *nonwoven*.

Woven geotextiles are made of *monofilament*, *multifilament*, or *fibrillated* yarns, or of slit films and tapes. The weaving process is of course very old, but nonwoven textile manufacture is a modern industrial development. Synthetic polymer fibers or filaments are continuously extruded and spun, blown or otherwise laid onto a moving belt. Then the mass of filaments or fibers are either *needle-punched*, in which the filaments are mechanically entangled by a series of small needles, or *heat bonded*, in which the fibers are in effect *welded* together by heat and/or pressure at their points of contact in the nonwoven mass.

Stiff geogrids with integral junctions are manufactured by extruding and orienting sheets of polypropylene or polyethylene. Flexible geogrids are made of polyester yams joined at the crossover points by knitting or weaving, and coated with a polymer.

For additional details on the composition, materials, and manufacturing of geomembranes and other geosynthetics, see Koerner (2005).

2.3 *Identification*

Geosynthetics are generically identified by: (1) polymer; (2) type of fiber or yarn, if appropriate; (3) type of geosynthetic; (4) mass per unit area or thickness, if appropriate; and (5) any additional information or physical properties necessary to describe the material. Four examples are:

- polypropylene staple fiber needle-punched nonwoven, 350 g/m^2;
- polyethylene net, 440 g/m^2 with 8 mm openings;
- polypropylene biaxial geogrid with 25 × 25 mm openings; and
- high-density polyethylene geomembrane, 1.5 mm thick.

3 Functions and applications

Geosynthetics have six primary functions: (1) filtration, (2) drainage, (3) separation, (4) reinforcement, (5) fluid barrier, and (6) protection.

Geosynthetic applications are usually defined by their primary, or principal, function. In a number of applications, in addition to the primary function, geosynthetics usually perform one or more secondary functions. It is important to consider both the primary and secondary functions in the design computations and specifications.

More than 150 separate applications of geosynthetics have been identified (Holtz et al., 1997; Koerner, 2005). A few examples are:

Geotextile filters replace graded granular filters in trench drains to prevent soils from migrating into drainage aggregate or pipes. They are also used as filters below riprap and other armor materials in coastal and riverbank protection systems. Geotextiles and geocomposites can also be used as drains, by allowing water to drain from or through soils of lower permeability. Examples include pavement edge drains, slope interceptor drains, and abutments and retaining wall drains.

Geotextiles are often used as separators to prevent fine-grained subgrade soils from being pumped into permeable, granular road bases and to prevent road base materials from penetrating into the underlying soft subgrade. Separators maintain the design thickness and roadway integrity.

Geogrid and geotextile reinforcement enables embankments to be constructed over very soft foundations. They are also used to construct stable slopes at much steeper

angles than would otherwise be possible, and an important reinforcement application is for retaining walls and abutments.

Geomembranes, thin-film geotextile composites, geosynthetic-clay liners, and field-coated geotextiles are used as fluid barriers to impede the flow of a liquid or gas from one location to another. This geosynthetic function has application in asphalt pavement overlays, encapsulation of swelling soils, and waste containment. In the sixth function, protection, the geosynthetic acts as a stress relief layer. A protective cushion of nonwoven geotextiles is often used to prevent puncture of geomembranes (by reducing point stresses) from stones in the adjacent soil or drainage aggregate during installation and while in service.

4 Design and selection

In the early days where there were only a few geotextiles available, design was mostly by trial and error and product selection was primarily by type or brand name. Today, however, with such a wide variety of geosynthetics available, this approach is inappropriate. When designing, selecting, and specifying geosynthetics, the recommended approach is the same as what is commonly done for any geotechnical engineering design. First, the design should be made without geosynthetics to see if they really are needed. If conventional solutions are impractical or uneconomical, then design calculations using reasonable engineering estimates of the required geosynthetic properties are carried out. Next, generic or performance type specifications are written so that the most appropriate and economical geosynthetic is selected, consistent with the properties required for its design functions, ability to survive construction, and its durability. In addition to conventional soils and materials testing, testing and properties evaluation of the geosynthetic is necessary. Finally, just as with any other construction, design with geosynthetics is not complete until construction has been satisfactorily carried out. Therefore, careful field inspection during construction is essential for a successful project. Additional discussion on all these points is given by Holtz et al. (1997).

5 Geosynthetics properties and tests

5.1 *Introduction*

Because of the wide variety of products available, with different polymers, filaments, weaving patterns or bonding mechanisms, thickness, mass, etc., geosynthetics have a considerable range of physical and mechanical properties. A further complicating factor is the variability of some properties, even within the same manufactured lot or roll; also, some differences may be due to the test procedures themselves.

Thus, determination of the design properties is not necessarily easy, although geosynthetic testing has progressed enormously in the last 25 yr. Standard procedures for testing geosynthetics for the determination of most properties have been developed by ASTM and other standards development organizations throughout the world, particularly in Europe, Japan, and Australia. The design properties required for a given design will depend on the specific application and the associated function(s) the geosynthetic is supposed to provide.

Geosynthetic properties can be classified as (1) general, (2) index, and (3) performance properties. See Holtz et al. (1997) for a listing of the various properties under

these categories, while Koerner and Hsuan (2001) describe test methods for the various geosynthetics properties, including those appropriate for remembrances and other products used for waste containment.

5.2 *General and index properties and tests*

General properties include the polymer, mass per unit area, thickness, roll dimensions and weight, specific gravity, etc. Because in most cases index tests do not give an actual design property, they do provide a qualitative assessment of the property of interest. When determined using standard test procedures, index test values can be used for product comparison, specifications, quality control purposes, and as an indicator of how the product might survive the construction process. These latter properties are called constructability or survivability properties.

Index tests include uniaxial mechanical strength (grab tensile; load-strain; creep, tear, and seam strength); multiaxial rupture strength (puncture, burst, and cutting resistance; flexibility); endurance or durability tests (abrasion resistance; UV stability; chemical and biological resistance; wet-dry and temperature stability); and hydraulic index tests (apparent opening size, percent open area; pore size distribution; porosity; permeability and permittivity; transmissivity).

5.3 *Performance properties and tests*

Performance properties require testing the geosynthetic and the soil together in order to obtain a direct assessment of the property of interest. Because performance tests should be conducted under design specific conditions and with soil samples from the site, these tests must be performed under the direction of the design engineer. Performance tests are not normally used in specifications; rather, geosynthetics should be preselected for performance testing based on index values, or performance test results should be correlated to index values for use in specifications. Examples of performance tests include in-soil stress-strain, creep, friction/adhesion, and dynamic tests; puncture; chemical resistance; and filtration or clogging resistance tests.

6 Specifications

Good specifications are essential for the success of any civil engineering project, and this is especially true for projects in which geosynthetics are to be used. Christopher and DiMaggio (1984) and Holtz et al. (1997) give guidance on writing generic and performance-based geotextile specifications. Generic specifications are just that, generic, and are based on the specific geosynthetic properties required for design, installation, and long-term performance. To specify a particular geosynthetic "brand name or its equivalent" can cause difficulties during installation. The contractor may select a product that has completely different properties than intended by the designer, and determination of what is "equivalent" is always a problem. Performance specifications utilize performance properties as described above.

Specifications based on "approved products" lists can also be developed if based on laboratory testing and experience with specific applications and conditions. Once an approved list has been established by an agency, new geosynthetics can be added after appropriate evaluation. Development of an approved list takes considerable initial

effort, but once established, it provides a simple and convenient method of specifying geosynthetics for routine applications.

All geosynthetic specifications should include the following items:

- general requirements
- specific geosynthetic properties
- seams and overlaps
- placement procedures
- repairs, and
- acceptance and rejection criteria.

General requirements include the types of geosynthetics, acceptable polymeric materials, and comments related to the stability and durability of the material. Geosynthetic manufacturers and representatives are good sources of information on these characteristics. Other items that should be specified in this section are instructions on storage and handling so products can be protected from exposure to ultraviolet light, dust, mud, or anything that may affect their performance. If pertinent, roll weight and dimensions may also be specified, and certification requirements should be included in this section.

Specific geosynthetic physical, index, and performance properties as required by the design must be listed. Properties should be given in terms of minimum (or maximum) average roll values (MARV) along with the appropriate test methods for those properties. MARVs are the smallest (or largest) anticipated average value that would be obtained for any roll tested (Holtz et al., 1997; Koerner, 2005). This average property value must exceed the minimum (or be less than the maximum) value specified for that property based on a particular test. Ordinarily it is possible to obtain a manufacturer's certification for MARVs, but it is also a good idea for the design engineer to have some verification testing performed for critical design and performance properties.

In virtually all geosynthetics applications, seams or overlaps are required and must be clearly specified. A minimum overlap of 0.3 m is recommended for all geotextile applications, but overlaps may be increased due to specific site and construction requirements. If overlaps will not work, then the geosynthetics must be seamed. Geotextiles are commonly seamed by sewing; see Holtz et al. (1997) for details. The specified seam strengths should equal the required strength of the geosynthetic, in the direction perpendicular to the seam length and using the same test procedures. (Seam strengths should not be specified as a percent of the geosynthetic strength.) Geogrids and geonets may be connected by mechanical fasteners, though the connection may be either structural or a construction aid (i.e., if strength perpendicular to the seam length is not required by design). Geomembranes are thermally or chemically bonded; see Koerner (2005) for details.

For sewn geotextiles, geomembranes, and structurally connected geogrids, the seaming material (thread, extrudate, or fastener) should consist of materials that have the same or greater durability as the geosynthetic being seamed. This is true for both factory and field seams.

Placement procedures should be specified in detail and in notes on the construction drawings. These procedures include grading and ground-clearing requirements, aggregate specifications, aggregate lift thickness, and equipment requirements. These

requirements are especially important if the geosynthetic was selected on the basis of survivability. Detailed placement procedures are given by Holtz et al. (1997).

Repair procedures for damaged sections of geosynthetics (i.e., rips and tears) should be detailed in the specifications.

Geosynthetic acceptance and rejection criteria should be clearly and concisely stated in the specifications. All installations should be observed by a competent inspector who is knowledgeable about placement procedures and design requirements. Sampling and testing requirements should also be specified.

7 Geosynthetics for soil reinforcement

The three primary soil reinforcement applications using geosynthetics are (1) reinforcing the base of embankments constructed on very soft foundations, (2) for soil slopes, increasing their stability and stable slope angle, and (3) reducing the earth pressures behind retaining walls and abutments. In the first two applications, geosynthetics permit construction that otherwise would be cost prohibitive or not technically feasible. In the case of retaining walls, significant cost savings are possible in comparison with conventional retaining wall construction. Other reinforcement and stabilization applications in which geosynthetics have also proven to be very effective include roads and railroads, large area stabilization, and natural slope reinforcement, but these applications are not discussed in this paper.

For each reinforcement application, the concept, design considerations, specific material properties, and construction considerations are discussed. In the case of retaining walls and abutments, because the wall face is such an important part of the structure, additional comments on wall facing systems are also given.

8 Reinforced embankments on soft foundations

8.1 *Concept*

The design and construction of embankments on soft foundation soils is a very challenging geotechnical problem. As noted by Leroueil and Rowe (2001), successful projects require a thorough subsurface investigation, properties determination, and settlement and stability analyses. If the settlements are too large or instability is likely, then some type of foundation soil improvement is warranted. Traditional soil improvement methods include preloading/surcharging with drains; lightweight fill; excavation and replacement; deep soil mixing, embankment piles, etc., as discussed by Holtz (1989) and Holtz et al. (2001b). Today, geosynthetic reinforcement must also be considered as a feasible foundation treatment alternative. In some situations, the most economical final design may be some combination of a traditional foundation treatment alternative together with geosynthetic reinforcement. Fig. 2a shows the basic concept for using geosynthetic reinforcement. Note that the reinforcement will not reduce the magnitude of long-term consolidation or secondary settlement of the embankment.

8.2 *Design considerations*

As with ordinary embankments on soft soils, the basic design approach for reinforced embankments is to design against failure. The ways in which embankments

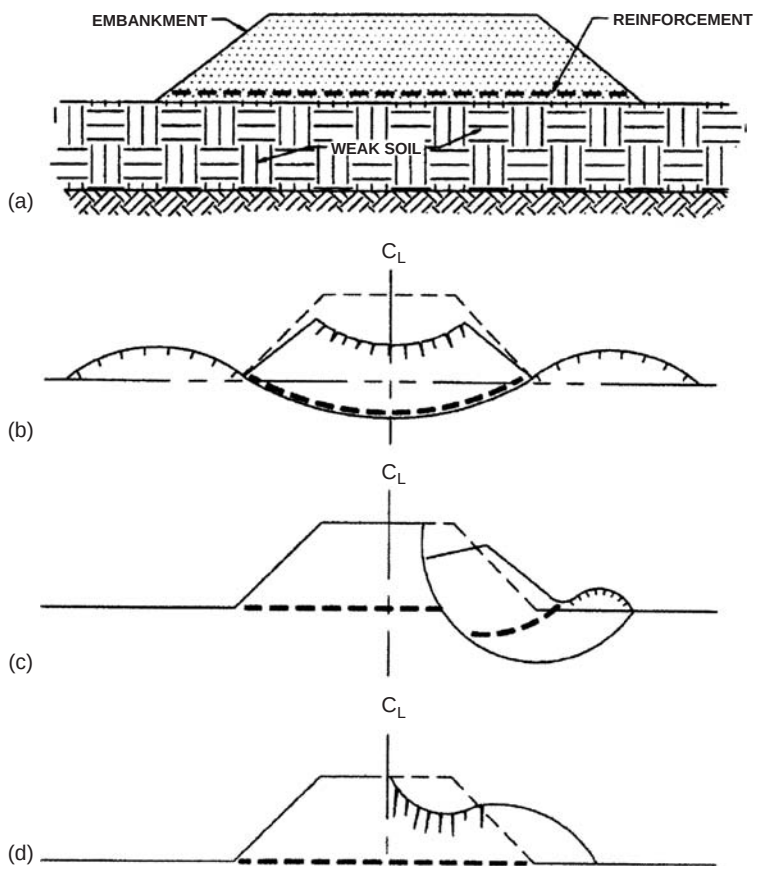


Figure 2 Reinforced embankments: a) concept; b) bearing failure; c) rotational failure; and d) lateral spreading. (After Bonaparte and Christopher, 1987, and Haliburton et al., 1978).

constructed on soft foundations can fail have been described by Terzaghi et al. (1996), among others. Fig. 2b-d shows unsatisfactory behavior that can occur in reinforced embankments. The three possible modes of failure indicate the types of stability analyses that are required for design. Overall bearing capacity of the embankment must be adequate, and the reinforcement should be strong enough to prevent rotational failures at the edge of the embankment. Lateral spreading failures can be prevented by the development of adequate shearing resistance between the base of the embankment, the reinforcement, and the foundation. In addition, an analysis to limit geosynthetic deformations must be performed. Finally, the geosynthetic strength requirements in the longitudinal direction, typically the transverse seam strength, must be determined. Discussion of these design concepts as well as detailed design procedures are given by Christopher and Holtz (1985), Bonaparte et al. (1987), Holtz (1989 and 1990), Humphrey and Rowe (1991), Holtz et al. (1997), and Leroueil and Rowe (2001).

The calculations required for stability and settlement utilize conventional geotechnical design procedures modified only for the presence of the reinforcement. Because the most critical condition for embankment stability is at the end of construction, the total stress method of analysis is usually performed, which is conservative since the analysis generally assumes that no strength gain occurs in the foundation soil. It is always possible of course to calculate stability in terms of effective stresses provided that effective stress shear strength parameters are available and an accurate estimate of the field pore pressures can be made during the project design phase. Because the prediction of in situ pore pressures in advance of construction is not easy, it is essential that the foundation be instrumented with high quality piezometers during construction to control the rate of embankment filling. Preloading and staged embankment construction are discussed in detail by Ladd (1991) and summarized by Leroueil and Rowe (2001).

8.3 *Material properties*

Based on the stability calculations, the minimum geosynthetic strengths required for stability at an appropriate factor of safety can be determined. In addition to its tensile and frictional properties, drainage requirements, construction conditions, and environmental factors must also be considered. Geosynthetic properties required for reinforcement applications are given in Table 1.

When properly designed and selected, high-strength geotextiles or geogrids can provide adequate embankment reinforcement. Both materials can be used equally well, provided they have the requisite design properties. There are some differences in how they are installed, especially with respect to seaming and field workability. Also, at some very soft sites, especially where there is no root mat or vegetative layer, geogrids

Table 1 Geosynthetic properties required for reinforcement applications.

Criteria and parameter	*Property*
Design requirements:	
Mechanical	
Tensile strength and modulus	Wide width strength and modulus
Seam strength	Wide width strength
Tension creep	Tension creep
Soil-geosynthetic friction	Soil-geosynthetic friction angle
Hydraulic	
Piping resistance	Apparent opening size
Permeability	Permeability and permittivity
Constructability requirements:	
Tensile strength	Grab strength
Puncture resistance	Puncture resistance
Tear resistance	Trapezoidal tear strength
Durability:	
UV stability (if exposed)	UV resistance
Chemical and biological (if required)	Chemical and biological resistance

may require a lightweight geotextile separator to provide filtration and prevent contamination of the embankment fill. However, a geotextile separator is not required if the fill can adequately filter the foundation soil.

A detailed discussion of geosynthetic properties and specifications is given by Holtz et al. (1997) and Koerner and Hsuan (2001), so only a few additional comments are given below.

The selection of appropriate fill materials is also an important aspect of the design. When possible, granular fill is preferred, especially for the first few lifts above the geosynthetic.

8.3.1 *Environmental considerations*

For most embankment reinforcement situations, geosynthetics have a high resistance to chemical and biological attack; therefore, chemical and biological compatibility is usually not a concern. However, in unusual situations such as very low (*i.e.*, <3) or very high (*i.e.*, >9) pH soils, or other unusual chemical environments (for example, industrial areas or near mine or other waste dumps), chemical compatibility with the polymer(s) in the geosynthetic should be checked. It is important to assure it will retain the design strength at least until the underlying subsoil is strong enough to support the structure without reinforcement.

8.3.2 *Constructability (survivability) requirements*

In addition to the design strength requirements, the geotextile or geogrid must also have sufficient strength to survive construction. If the geosynthetic is ripped, punctured, torn or otherwise damaged during construction, its strength will be reduced and failure could result. Constructability property requirements are listed in Table 1. (These are also called survivability requirements.)

See Christopher and Holtz (1985) and Holtz et al. (1997) for specific property requirements for reinforced embankment construction with varying subgrade conditions, construction equipment, and lift thicknesses. For all critical applications, high to very high survivability geotextiles and geogrids are recommended.

8.3.3 *Stiffness and workability*

For extremely soft soil conditions, geosynthetic workability will be an important consideration. Workability is the ability of a geosynthetic to support workpersons during initial placement and seaming operations as well as to support construction equipment during the first lift placement. Workability depends on the stiffness of the geosynthetic; however, stiffness evaluation techniques apparently do not correlate with field workability (Tan, 1990). See Holtz et al. (1997) for recommendations on stiffness.

8.4 *Construction*

The importance of proper construction procedures for geosynthetic reinforced embankments cannot be overemphasized. A specific construction sequence is usually required in order to avoid failures during construction. Appropriate site preparation, low ground pressure equipment, and small initial lift thickness are very important, and reduced loads of hauling vehicles may be required. Clean granular fill is recommended,

especially for the first few construction lifts, and proper fill placement, spreading, and compaction procedures are required. A detailed discussion of construction procedures for reinforced embankments on very soft foundations is given by Christopher and Holtz (1985) and Holtz et al. (1997).

It should be noted that all geosynthetic seams must be positively joined; overlaps will simply not be sufficient. For geotextiles, this means sewing; for geogrids, some type of positive clamping arrangement must be used. Inspection is essential, as the seams are the "weak link" in the system, and seam failures are quite common in improperly constructed embankments. Finally, soft ground construction projects usually require geotechnical instrumentation for proper control of construction and fill placement; see Holtz (1989) and Holtz et al. (200la) for recommendations.

9 Reinforced steep slopes

9.1 *Concept*

The first use of geosynthetics for the stabilization of steep slopes was for the reinstatement of failed slopes. Cost savings resulted because the slide debris could be reused in the repaired slope (together with geosynthetic reinforcement), rather than importing select materials to reconstruct the slope. Even if foundation conditions are satisfactory, costs of fill and right-of-way plus other considerations may require a steeper slope than is stable in compacted embankment soils without reinforcement. As shown in Fig. 3, multiple layers of geogrids or geotextiles may be placed in a fill slope during construction or reconstruction to reinforce the soil and provide increased slope stability. Most steep slope reinforcement projects are for the construction of new embankments, alternatives to retaining walls, widening of existing embankments, buttresses for unstable natural slopes, and repair of failed slopes.

Another use of geosynthetics in slopes is for compaction aids (Fig. 3). In this application, narrow geosynthetic strips, 1 to 2 m wide, are placed at the edge of the fill slope to provide increased lateral confinement at the slope face, and therefore increased compacted density over that normally achieved. Even modest amounts of reinforcement in compacted slopes have been found to prevent sloughing and reduce slope erosion. In some cases, thick nonwoven geotextiles with in-plane drainage capabilities allow for rapid pore pressure dissipation in compacted cohesive fill soils.

9.2 *Design considerations*

The overall design requirements for reinforced slopes are similar to those for unreinforced slopes—the factor of safety must be adequate for both the short- and long-term conditions and for all possible modes of failure. These include: (1) internal, where the failure plane passes through the reinforcing elements; (2) external, where the failure surface passes behind and underneath the reinforced mass; and (3) compound, where the failure surface passes both behind and through the reinforced soil mass.

Reinforced slopes are analyzed using modified versions of classical limit equilibrium slope stability methods (e.g., Terzaghi et al., 1996). Potential circular or wedge-type failure surfaces are assumed, and the relationship between driving and resisting forces

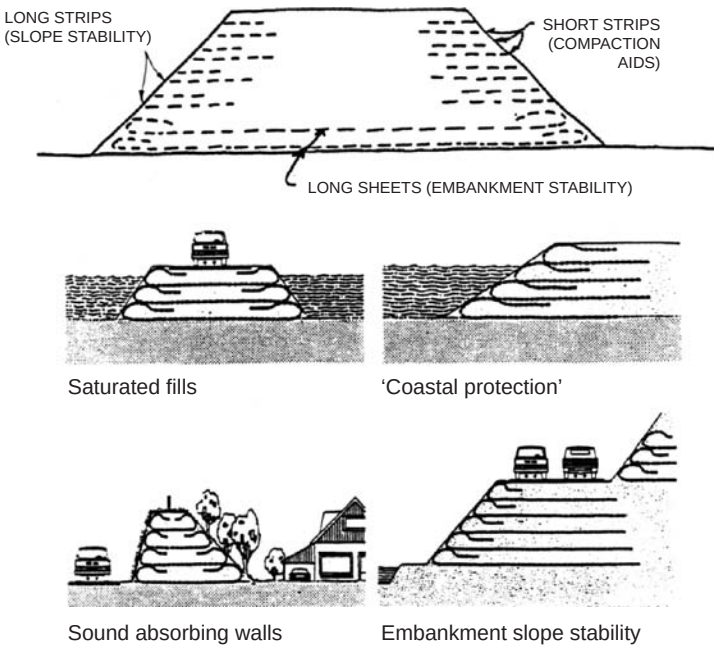


Figure 3 Examples of multilayer geosynthetic slope reinforcement.

or moments determines the factor of safety. Based on their tensile capacity and orientation, reinforcement layers intersecting the potential failure surface increase the resisting moment or force. The tensile capacity of a reinforcement layer is the minimum of its allowable pullout resistance behind, or in front of, the potential failure surface and/or its long-term design tensile strength, whichever is smaller. A variety of potential failure surfaces must be considered, including deep-seated surfaces through or behind the reinforced zone, and the critical surface requiring the maximum amount reinforcement determines the slope factor of safety. The reinforcement layout and spacing may be varied to achieve an optimum design. Computer programs are available for reinforced slope design which include searching routines to help locate critical surfaces and appropriate consideration of reinforcement strength and pullout capacity.

Additional information on reinforced slope design is available in Christopher et al. (1990), Christopher and Leshchinsky (1991), Berg (1993), Holtz et al. (1997), and Bathurst and Jones (2001).

For landslide repairs or when used as a buttress for natural soil slopes, it is very important that the cause of original failure or the cause of the potential instability be determined and addressed in order to insure that the new reinforced soil slope will not have the same problems. Particular attention must be paid to drainage. Just as with natural slopes, infiltrating ground or surface water can cause instability in reinforced slopes. It is also necessary to identify any weak seams that could affect stability. Holtz et al. (2001b) present a case history of the failure during an earthquake of a steep

reinforced soil buttress that also had potential static instability due to a weak seam in the natural slope as well as poor drainage of the backfill.

9.3 *Material properties*

Geosynthetic properties required for reinforced slopes are similar to those listed in Table 1. Properties are required for design (stability), constructability, and durability. Allowable tensile strength and soil-geosynthetic friction are most important for stability design. Because of uncertainties in creep strength, chemical and biological degradation effects, installation damage, and joints and connections, a partial factor or reduction factor concept is recommended. The ultimate wide width strength is reduced for these various factors, and the reduction depends on how much information is available about the geosynthetics at the time of design and selection. Berg (1993), Holtz et al. (1997), and Koerner and Hsuan (2001) give details about the determination of the allowable geosynthetic tensile strength. They also describe how soil-geosynthetic friction is measured or estimated.

An inherent advantage of geosynthetic reinforcement is their longevity, especially in normal soil environments. Recent studies have indicated that the anticipated half-life of reinforcement geosynthetics in between 500 and 5000 yr in typical soil environments; in unusual environmental conditions, strength characteristics can be adjusted to account for potential degradation.

Any soil suitable for embankment construction can be used in a reinforced slope system. From a reinforcement point of view alone, even lower-quality soil than conventionally used in unreinforced slope construction may be used. However, higher-quality materials offer less durability concerns, are easier to place and compact, which tends to speed up construction, and they have fewer problems with drainage. See Berg (1993) and Holtz et al. (1997) for discussion of soil gradation, compaction, unit weight, shear strength, and chemical composition.

9.4 *Construction*

Similarly to reinforced embankments, proper construction is very important to insure adequate performance of a reinforced slope. Considerations of site preparation, reinforcement and fill placement, compaction control, face construction, and field inspection are given by Berg (1993) and Holtz et al. (1997).

10 Reinforced retaining walls and abutments

10.1 *Concept*

Retaining walls are required where a soil slope is uneconomical or not technically feasible. When compared with conventional retaining structures, walls with reinforced soil backfills have several significant advantages. They are very cost effective, especially for higher walls. Furthermore, these systems are more flexible than conventional earth retaining walls such as reinforced concrete cantilever or gravity walls. Therefore, they are very suitable for sites with poor foundations and for seismically active areas.

Modern reinforced soil walls were developed in the mid 1960s by the French architect, H. Vidal; Vidal called his system “Reinforced Earth”. As shown in Fig. 4,

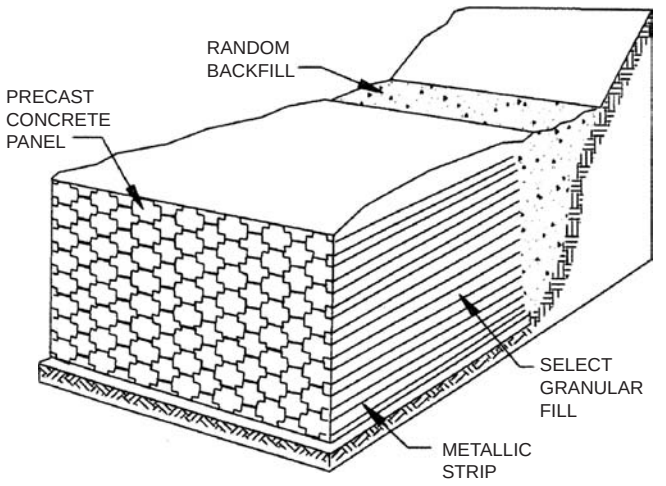


Figure 4 Component parts of a Reinforced Earth wall.

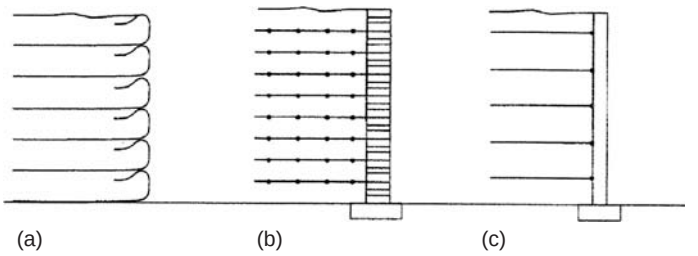


Figure 5 Reinforced retaining wall systems using geosynthetics; (a) with wrap-around geosynthetic facing, (b) with segmental or modular concrete block, and (c) with full-height (propped) precast panels.

metallic (galvanized steel) straps are attached to precast concrete panels and incorporated into compacted granular soil backfill. The design and construction of Vidal-type reinforced earth walls are now well established, and many thousands have been successfully built throughout the world in the last 35 yr. Other similar proprietary reinforcing systems have also been developed using steel bar mats, grids, and gabions.

The use of geotextiles as reinforcing elements started in the early 1970's because of concern over possible corrosion of metallic reinforcement. Systems using sheets of geosynthetics rather than steel strips are shown in Fig. 5.

Among the significant benefits of both steel and geosynthetic reinforced soil walls are the wide variety of facing systems available to designers and owners. Articulated concrete panels and blocks are probably most common, although wire gabions, timber beams, shotcrete, and cast-in-place concrete have also been used.

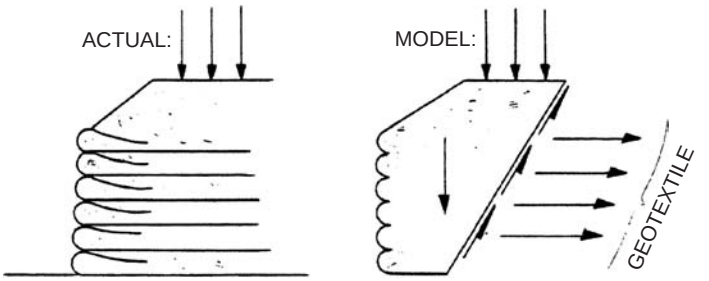


Figure 6 Actual geosynthetic-reinforced wall compared to its analytical model.

Walls reinforced with steel have been successfully built to heights of more than 40 m, whereas the maximum heights so far of geosynthetic-reinforced walls are less than 20 m.

10.2 *Design considerations*

Except for the added complexity of the internal stability of the reinforced section, the design of reinforced walls is very similar to conventional retaining wall design practice. External stability is calculated in the conventional way – the bearing capacity must be adequate, the reinforced section may not slide or overturn, and overall slope stability must be adequate. Surcharges (live and dead loads; distributed and point loads) are considered in the conventional manner. Settlement of the reinforced section also should be checked if the foundation is compressible.

A number of different approaches to internal design of geotextile-reinforced retaining walls have been proposed (Christopher et al., 1990; Allen and Holtz; 1991; Holtz, 1995), but the oldest and most common-and most conservative–method is the tieback wedge analysis. It utilizes classical earth pressure theory combined with tensile resisting "tiebacks" that extend back of the assumed failure plane (Fig. 6). The K_A (or K_o) is assumed, depending on the stiffness of the facing and the amount of yielding likely to occur during construction, and the earth pressure at each vertical section of the wall is calculated. This earth pressure must be resisted by the geosynthetic reinforcement at that section.

To design against failure of the reinforcement, there are two possible limiting or failure conditions: rupture of the geosynthetic and pullout of the geosynthetic. The corresponding reinforcement properties are the tensile strength of the geosynthetic and its pullout resistance. In the latter case, the geosynthetic reinforcement must extend some distance behind the assumed failure wedge so that it will not pull out of the backfill. Typically, sliding of the entire reinforced mass controls the length of the reinforcing elements. For a detailed description of the tieback wedge method, see Christopher and Holtz (1985), Bonaparte et al. (1987), Allen and Holtz (1991), and Holtz et al. (1997). Recent research (e.g., Lee et al., 1999; Lee, 2000; Bathurst et al., 2000) has indicated that the tieback wedge approach is overly conservative and uneconomical, and modifications and deformation-based designs are rapidly being developed (See, e.g., Allen et al., 2003 and 2004).

Other important internal design considerations include backfill drainage and potential seismic loading.

In terms of its behavior, fundamentally a reinforced wall is just a very steep reinforced soil slope. Yet the approach to the analysis and design of reinforced soil walls is based on classical earth pressure theory, while stability analyses of reinforced soil slopes are similar to traditional soil mechanics slope stability analyses.

10.3 *Material properties*

Geosynthetic properties required for reinforced walls are similar to those listed in Table 1 and discussed in Section 9.3 for reinforced slopes. Properties are required for design (stability), constructability, and durability. Allowable tensile strength and soil-geosynthetic friction are required for stability design, and similar to reinforced slopes, a partial factor or reduction factor approach is common. The ultimate wide width strength is reduced to account for uncertainties in creep strength, chemical and biological degradation effects, installation damage, and joints and connections. Berg (1993), Holtz et al. (1997), Elias et al. (2001), and Koerner and Hsuan (2001) give details about the determination of the allowable geosynthetic tensile strength. They also describe how soil-geosynthetic friction is measured or estimated.

The discussion on durability and longevity of geosynthetic reinforcement in Section 9.3 is pertinent for reinforced walls as well.

Soil properties required include gradation, percent fines, chemical composition, compaction, unit weight, and shear strength. To insure stability, appropriate consideration of the foundation and overall slope stability at the site is also important (Holtz et al., 2001b).

Backfill for geosynthetic reinforced walls should be free draining if at all possible. If it is not free draining, then adequate internal drainage for infiltrating surface and/or groundwater must be provided in the design. This is important for internal and face stability because drainage outward through the wall facing system may not be adequate.

10.4 *Wall facing considerations*

A significant advantage of geosynthetic reinforced walls over conventional retaining structures is the variety of facings that can be used and the resulting aesthetic options that can be provided. Aesthetic requirements often determine the type of facing systems. Anticipated deflection of the wall face, both laterally and downward, may place further limitations on the type of facing system selected. Tight construction specifications and quality inspection are necessary to insure that the wall face is constructed properly; otherwise an unattractive wall face, or a wall face failure, could result.

Facing systems can be installed (1) as the wall is being constructed, or (2) after the wall is completed. Facings installed as the wall is constructed include segmental and full height pre-cast concrete panels, interlocking pre-cast concrete blocks, welded wire panels, gabion baskets, treated timber facings, and geosynthetic face wraps. In these cases, the geosynthetic reinforcement is attached directly to the facing element. Systems installed after construction include shotcrete, cast-in-place concrete fascia, and pre-cast concrete or timber panels; the panels are attached to brackets placed between the layers of the geosynthetic wrapped wall face at the end of wall construction or

after wall movements are complete. Facings constructed as the wall is constructed must either allow the geosynthetic to deform freely during construction without any buildup of stress on the face, or the facing connection must be designed to take the stress. Although most wall design methods assume that the stress at the face is equal to the maximum horizontal stress in the reinforced backfill, measurements show that considerable stress reduction occurs near the face, depending on the flexibility of the face. See Allen and Holtz (1991) and Holtz et al. (1997) for a detailed discussion of wall facing systems.

10.5 *Construction*

Construction procedures for geosynthetic reinforced walls and abutments are relatively simple and straightforward, and details are given by Christopher and Holtz (1985), Holtz et al. (1997), Collin (2001), and Elias et al. (2001). Basically, the backfill material is placed and compacted in layers, with the geosynthetic reinforcement placed between each compacted lift or often between multiple lifts, depending on the specific design requirements. In either case, the soil and reinforcement are supported at the face either by a temporary support system or by the permanent facing elements themselves.

For a geotextile or geogrid wrapped face wall, a temporary wood form can be used. Sometimes, a stiff welded wire mesh together with a geotextile liner permanently retains the backfill at the wall face. If precast concrete elements or modular blocks are used for the wall face, then support is provided by the permanent concrete face element. In this case, the reinforcement is connected to the face element either by friction between adjacent blocks by some type of positive connection such as pins, concrete grooves, or the shape of the concrete elements.

Especially with critical structures such as GRS walls, competent and professional inspection of the wall construction is absolutely essential for a successful project. The design engineer should develop procedures to ensure that the specified materials are delivered to the project site, that the geosynthetic is not damaged during installation, and that the specified sequence of construction operations is explicitly followed. Inspectors should be very familiar with all details of the material specifications and construction, including backfill and geosynthetic placement, drainage provisions, and wall face construction and alignment. Depending on the experience of the contractor and field inspectors, preconstruction review meetings may be appropriate.

For additional information about GRS construction inspection, see Holtz et al. (1997) and Elias et al. (2001). The latter reference also gives suggestions for inspection of prefabricated concrete elements and joint materials, if used, along with additional details on construction control and performance monitoring.

10.6 *Avoiding failures*

No owner, designer, or contractor likes failures, especially on their own projects. However, in spite of all good intentions, failures of GRS walls are surprisingly common, especially with proprietary precast segmental concrete block-faced wall systems. Because current design procedures and geosynthetic properties' specifications are overly conservative, it is almost impossible for a GRS wall to fail due to internal overstress or pullout of the geosynthetic reinforcement. So why are there so many

failures? It appears that most failures are due to (1) inadequate geotechnical design, particularly of the foundation and back slope of the wall, (2) inadequate internal or external drainage, or (3) problems in construction. The latter include poor inspection and quality control, poor compaction, use of inappropriate backfill materials, contamination of drainage materials or their elimination, and lack of attention to facing connections and details.

In some jurisdictions, walls are classified as structures, and structural engineers are legally required to be responsible of the wall design, a ludicrous situation with GRS walls that obtain their stability with soil and geosynthetics. Often in site development work, the lines of responsibility between wall designers, material suppliers, site civil engineers, construction inspectors, contractors and subcontractors, and owners are often very fuzzy and lead to unwanted performance or even catastrophic collapses of GRS walls. Attention to all the details of design, materials, and construction mentioned above will insure successful GRS wall projects.

Acknowledgments

This paper is based on research sponsored by the U.S. National Science Foundation, U.S. Federal Highway Administration, U.S. Air Force, the Indiana Department of Highways, the Washington State Department of Transportation, and several geosynthetics manufacturers. T.M. Allen, J.R. Bell, R.R. Berg, S.R. Boyle, R. Bonaparte, E. Boutrup, B.R. Christopher, J.A. DiMaggio, J.P. Giroud, T.H. Hatzigogos, D.N. Humphrey, W.F. Lee, R.M. Koerner, F. Saidin, and O.Wager influenced my thinking and helped me with various aspects of this work. I very much appreciate all this support and assistance.

References

Allen, T.M. and Holtz, R.D. (1991) "Design of Retaining Walls Reinforced with Geosynthetics", State of the Art Paper, session 8A on Earth Reinforcement, *Geotechnical Engineering Congress 1991*, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 970–987.

Allen, T.M., Bathurst, R.J., Holtz, R.D., Walters, D. and Lee, W.F. (2003) "A New Working Stress Method for Prediction of Reinforcement Loads in Geosynthetic Walls", *Canadian Geotechnical Journal*, Vol. 40, No. 5, pp. 976–994.

Allen, T.M., Bathurst, R.J., Lee, W.F., Holtz, R.D. and Walters, D. (2004) "A New Working Stress Method for Prediction of Reinforcement Loads in Steel Reinforced MSE Walls", *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 130, No. 11, pp. 1109–1120.

Bathurst, R.J. and Jones, C.J.F.P. (2001) Chapter 17 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 501–537.

Bathurst, R.J., Walters, D., Vlachopoulos, N., Burgess, P. and Allen, T.M. (2000) "Full Scale Testing of Geosynthetic Reinforced Walls", *Advances in Transportation and Geoenvironinental Systems Using Geosynthetics*, Proceedings of sessions at GeoDenver 2000, J.G. Zornberg and B.R. Christopher, Editors, Geotechnical Special Publication No. 103, ASCE, pp. 201–217.

Berg, R.R. (1993) *Guidelines for Design, Specification, & Contracting of Geosynthetic Mechanically Stabilized Earth Slopes on Firm Foundations*, Federal Highway Administration, Report No. FHWA-SA-93-025, 87 pp.

Bonaparte, R. and Christopher, B.R. (1987) "Design and Construction of Reinforced Embankments over Weak Foundations", *Transportation Research Record 1153*, TRB, pp. 226–39.

Bonaparte, R., Holtz, R.D. and Giroud, J.P. (1987) "Soil Reinforcement Design Using Geotextiles and Geogrids", in *Geotextile Testing and the Design Engineer, ASTM Special Technical Publication 952*, J.E. Fluet, editor, American Society for Testing and Materials, pp. 69–116.

Christopher, B.R. and Di Maggio, J.A. (1984) "Specifying Geotextiles", *Geotechnical Fabrics Report*, Vol. 2, No. 2, pp. 21–25.

Christopher, B.R., Gill, S.A., Giroud, J.P., Juran, I., Mitchell, J.K., Schiosser, F. and Dunnicliff, J. (1990) *Reinforced Soil Structures*, Vol. I: Design and Construction Guidelines, Federal Highway Administration, Report No. FHWA-RD-89-043, 287 pp.

Christopher, B.R. and Holtz, R.D. (1985), *Geotextile Engineering Manual*, U.S. Federal Highway Administration, Washington, D.C., FHWA-TS-86/203, 1044 pp.

Christopher, B.R. and Leshchinsky, D. (1991) "Design of Geosynthetically Reinforced Slopes", State of the Art Paper, Session 8A on Earth Reinforcement, *Geotechnical Engineering Congress 1991*, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 998–1005.

Collin, J.G., Editor (2001) *Design Manual for Segmental Retaining Walls*, Second Edition, National Concrete Masonry Association, NCMA Publication No. TR 127A, 289 pp.

Elias, V., Christopher, B.R. and Berg, R.R. (2001) *Mechanically Stabilized Earth Walls and Reinforced Soil Slopes–Design and Construction Guidelines*, U.S. Federal Highway Administration, National Highway Institute, Washington, D.C., Publication No. FHWA NHI-00-043, 394 pp.

Haliburton, T.A., Anglin, C.C. and Lawmaster, J.D. (1978) *Selection of Geotechnical Fabrics for Embankment Reinforcement*, Report to US Army Engineer District, Mobil, Oklahoma State University, Stillwater, 138 pp.

Holtz, R.D. (1989) "Treatment of Problem Foundations for Highway Embankments", *Synthesis of Highway Practice 147*, National Cooperative Highway Research Program, Transportation Research Board, 72 pp.

Holtz, R.D. (1990) "Design and Construction of Geosynthetically Reinforced Embankments on Very Soft Soil", State of the Art Paper, Session 5, *Performance of Reinforced Soil Structures*, Proceedings of the International Reinforced Soil Conference, British Geotechnical Society, Glasgow, Scotland, A. McGown, K. Yeo, and K.Z. Andrawes, Editors, Th. Telford (London), pp. 391–402.

Holtz, R.D. (1995) "Retaining Walls Reinforced with Geosynthetics: From Broms (1977, 1978) to the Present", *Proceedings* of the Bengt B. Broms Symposium on Geotechnical Engineering, Singapore, World Scientific, pp. 181–194.

Holtz, R.D., Christopher, B.R. and Berg, R.R. (1997) *Geosynthetic Engineering*, BiTech Publishers, Vancouver, British Columbia, Canada., 451 pp.

Holtz, R.D., Kramer, S.L., Hsieh, C.W., Huang, A.B. and Lee, W.F. (2001a) "Failure Analysis of an 80 m High Geogrid Reinforced Wall in Taiwan", *Proceedings* of the 15th International Conference on Soil Mechanics and Geotechnical Engineering, Istanbul, Turkey, Volume 2, pp. 1159–1162.

Holtz, R.D., Shang, J.Q. and Bergado, D.T. (2001b) "Soil Improvement", Chapter 15 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 429–462.

Humphrey, D.N. and Rowe, R.K. (1991) "Design of Reinforced Embankments—Recent Developments in the State-of-the-Art", State of the Art Paper, Session 8A on Earth Reinforcement, *Geotechnical Engineering Congress 1991*, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 1006–1020.

Koerner, R.M. (2005) *Designing with Geosynthetics*, 5th ed., Prentice-Hall, 706 pp.

Koerner, R.M. and Hsuan, Y.G. (2001) "Geosynthetics: Characteristics and Testing", Chapter 7 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 173–196.

Ladd, C.C. (1991) "Stability Evaluation During Staged Construction", 22nd Terzaghi Lecture, *Journal of Geotechnical Engineering*, ASCE, Vol. 117, No. 4, pp. 537–615.

Lee, W.F. (2000) "Internal Stability Analysis of Geosynthetic Reinforced Retaining Walls", Ph.D. dissertation, University of Washington, Seattle, 380 pp.

Lee, W.F., Holtz, R.D. and Allen, T.M. (1999) "Full Scale Geosynthetic Reinforced Retaining

Leroueil, S. and Rowe, R.K. (2001) "Embankments over Soft Soil and Peat", Chapter 16 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 463–499.

Tan, S.L. (1990) "Stress-Deflection Characteristics of Soil Soils Overlain with Geosynthetics", MSCE Thesis, University of Washington, Seattle, 146 pp.

Terzaghi, K., Peck, R.B. and Mesri, G. (1996) *Soil Mechanics in Engineering Practice*, 3rd Edition, Wiley, 549 pp.

Walls: A Numerical Parametric Study", *Proceedings* of Geosynthetics '99, Boston, Massachusetts, Industrial Fabrics Association International, Vol. 2, pp. 935–948.

Chapter 7

Risk-based decision making for energy infrastructure

R.B. Gilbert

The University of Texas at Austin, Austin, TX, USA

ABSTRACT: The infrastructure that supports the production, processing and distribution of energy is vital for a stable and healthy world. The objective of this paper is to illustrate the application of a risk-based framework to support the important decisions required in developing and maintaining the energy infrastructure. A series of case history examples is presented to demonstrate how this framework can be applied in practice. The value of the framework is to provide a context and a perspective, both qualitative and quantitative, for balancing costs and benefits in managing risks.

1 Introduction

The infrastructure that supports the production, processing and distribution of energy is vital for a stable and healthy world. The benefits of an economical, safe and sustainable supply of energy are unlimited. However, the costs associated with developing and maintaining this infrastructure are enormous. In addition, the potential for harm to the economy, human health and the environment is significant if the infrastructure does not function as intended. Finally, there is substantial uncertainty in the performance of this infrastructure, particularly considering the numerous hazards that threaten its integrity.

The objective of this paper is to illustrate the application of a risk-based framework to support the important decisions required in developing and maintaining infrastructure for energy. This risk-based framework is founded on fundamental principles of decision analysis and probability theory. A series of case history examples is presented to demonstrate how these principles can be applied in practice to help in answering the following questions: What factor of safety should be used in design? Is more information needed to make a decision? How can different types of information be combined in making a decision? How can risk be assessed? How does risk impact public policy? How best to communicate risk?

2 What factor of safety should be used?

A very common question is what factor of safety to use for design. The simple answer is to look it up in a design code or guidance document. However, design codes and guidance documents cannot possibly cover the range of possibilities in design, particularly as we continue to develop new technologies and encounter conditions never before

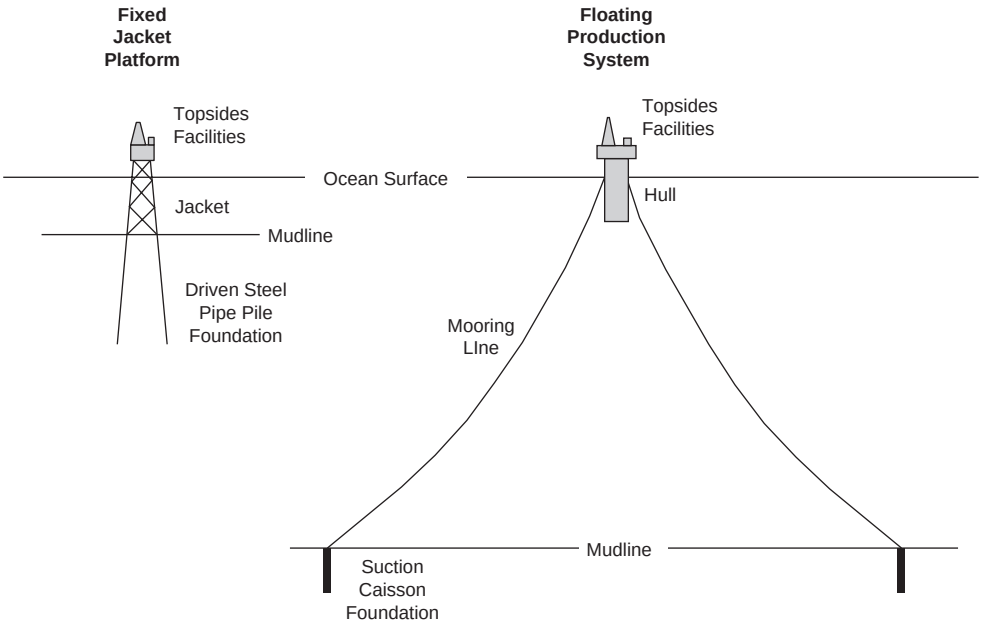


Figure 1 Schematic of offshore production systems (from Gilbert et al. 2008).

imagined. The following examples of designing foundations for production facilities illustrate the value in adopting a risk-based perspective in design.

2.1 *Fixed offshore production system*

A fixed jacket structure for gas production is to be located in about 100 m of water (Fig. 1). The soils at the site are normally consolidated marine clays inter-bedded with layers every 30 m or so of overconsolidated clay and channels of dense sand. The design for the pile foundations was governed by their axial capacity; they were designed to be 120-m long. The challenge is that the piles could not be driven to their final penetration because they became stuck in a dense sand layer at a depth of about 105 m below the mudline.

The decision is whether to leave the piles short or advance them further. A decision tree to frame this decision is shown in Fig. 2 (see Benjamin and Cornell 1970 and Ang and Tang 1984 for details on formal decision analysis). There is uncertainty about whether the foundation will be adequate or not, even if the piles are driven to their design length. An inadequate foundation means that the structure will be toppled in a hurricane sometime during its design life. The probabilities that the foundation will be inadequate (or fail) if the piles are left short, $p_{F,short}$, and driven further, $p_{F,long}$, reflect this uncertainty in performance. The cost of driving the piles further includes the cost of bringing drilling equipment and a larger pile driving hammer to this remote location.

A rational basis for making a decision is to compare the expected cost, which is the sum of all possible costs multiplied by their respective probabilities. The two alternatives have the following expected costs:

$$\begin{aligned} E(\text{Cost}_{\text{short}}) &= 100 p_{\text{F, short}} \ (\$\text{MM}) \\ E(\text{Cost}_{\text{long}}) &= 2 + 100 p_{\text{F, long}} \ (\$\text{MM}) \end{aligned} \tag{1}$$

The input needed to solve this problem is the probability of failure for each alternative.

These probabilities of failure can be quantified using the following simplified model:

$$\begin{aligned} p_F &= P(\text{Load} > \text{Capacity}) \\ &\cong \Phi\left(-\frac{\ln(FS_{\text{median}})}{\sqrt{\delta_{\text{load}}^2 + \delta_{\text{capacity}}^2}}\right) \end{aligned} \tag{2}$$

where P(Load > Capacity) is the probability that the axial load exceeds the capacity at some point in the design life, which is also referred to as the lifetime probability of failure; FS_{median} is the median factor of safety, which is defined as the ratio of the median capacity to the median load; and δ is the coefficient of variation (c.o.v.), which is defined as the standard deviation divided by the mean value for that variable. Equation (2) assumes that the load and capacity each follow independent lognormal distributions, a reasonable assumption in typical reliability analyses for offshore foundations (e.g., Tang and Gilbert 1993).

The median factor of safety in Eq. (2) is related as follows to the design factor of safety:

$$FS_{\text{median}} = FS_{\text{design}} \times \frac{\left(\frac{\text{capacity}_{\text{median}}}{\text{capacity}_{\text{design}}}\right)}{\left(\frac{\text{load}_{\text{median}}}{\text{load}_{\text{design}}}\right)} \tag{3}$$

where the subscript "design" indicates the value used to design the foundation (also referred to as the "nominal" value). In a load and resistance factor design recipe, FS_{design} in Eq. (3) would be replaced by the ratio of the load factor to the resistance factor (the safety margin). The ratios of the median to design values represent biases between the median or most likely value in the design life and the value that is used in the design check. For context, the median factor of safety is generally between three and five for a pile in a typical jacket platform.

The c.o.v. values Eq. (2) represent uncertainty in the load and the capacity. For an offshore foundation, the uncertainty in the load is generally due to variations in the occurrence and strength of hurricanes at the platform site over the design life. The uncertainty in the capacity is due primarily to variations between the actual capacity

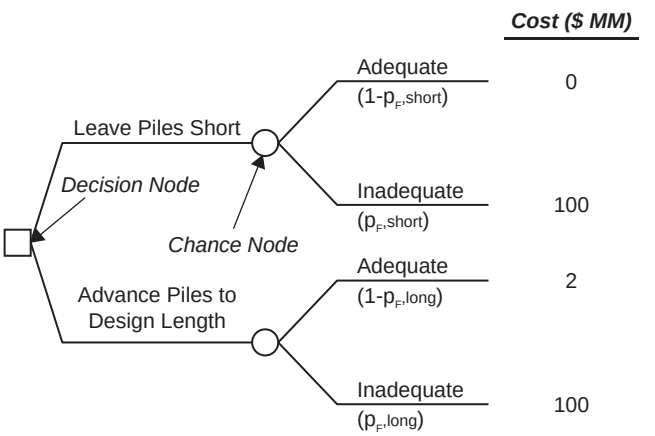


Figure 2 Decision tree for example with pile driving difficulty.

in a storm load compared to the capacity predicted using the design method. The denominator in Eq. (2) is referred to as the total coefficient of variation:

$$\delta_{\text{total}} = \sqrt{\delta_{\text{load}}^2 + \delta_{\text{capacity}}^2} \tag{4}$$

Typical c.o.v. values for a jacket platform range from 0.3 to 0.5 for the load, 0.3 to 0.5 for the capacity, and 0.4 to 0.7 for the total.

The relationship between the probability of failure and the median factor of safety and total c.o.v. is shown on Fig. 3. An increase in the median factor of safety and a decrease in the total c.o.v. both reduce the probability of failure. The calculated failure probability for a single pile in a typical jacket foundation over a 20-year design life is between 0.005 and 0.05 (Fig. 3). Note that the event of foundation failure, i.e. axial overload of a single pile in the foundation, does not necessarily lead to collapse of a jacket. Failure probabilities for the foundation system are about ten times smaller than those for a single pile (Tang and Gilbert 1993). Therefore, the probability of failure for the foundation system in a typical jacket is on the order of 0.0005 to 0.005.

For the example in Fig. 2, the piles were designed according to the following information:

$\text{load}_{\text{design}} = 13\,\text{MN}$
$\text{FS}_{\text{design}} = 1.5$
$\text{capacity}_{\text{design, long}} = 19.5\,\text{MN}$
$\text{load}_{\text{median}}/\text{load}_{\text{design}} = 1.5$
$\text{capacity}_{\text{median}}/\text{capacity}_{\text{design, long}} = 1.3$
$\text{FS}_{\text{median, long}} = 2.9$
$\delta_{\text{load}} = 0.3$
$\delta_{\text{capacity, long}} = 0.3$
$\delta_{\text{total, long}} = 0.42$

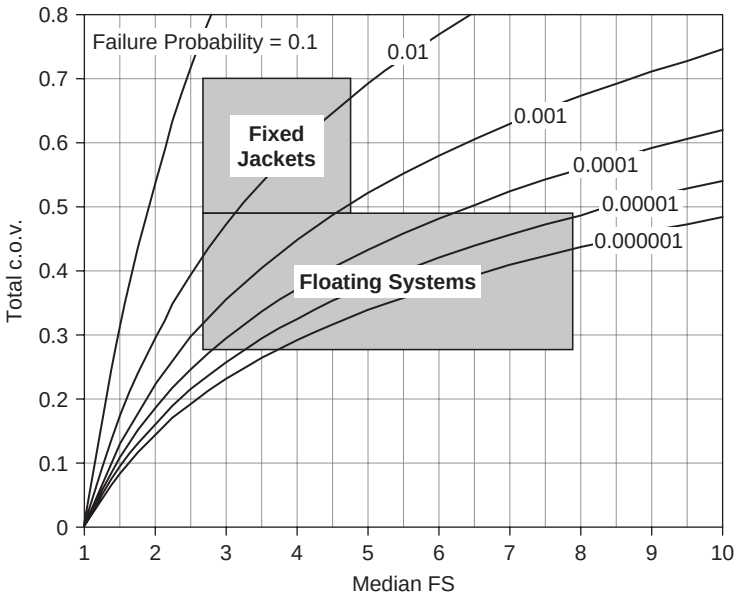


Figure 3 Simplified reliability model for offshore foundations (from Gilbert et al. 2008).

Hence, the probability of failure for an individual pile, driven to the design penetration, is 0.006 from Eq. (2) or Fig. 3. Accounting for system redundancy, the probability of failure for the as-designed foundation system is $p_{F,\,long} \cong 0.006/10 = 0.0006$ and the expected cost associated with the as-designed piles is $E(C_{long}) = \$2.06$ MM from Eq. (2).

The expected costs for the two alternatives are compared in Fig. 4 as a function of the probability of failure for the foundation system if the piles are left short. If the probability of failure for the foundation system with the short piles is less than 0.004, then this alternative will have the lower expected cost and be preferred (Fig. 4). In terms of a single pile, the target probability of failure for the short piles is about 0.04.

Do the short piles achieve a level of reliability such that the probability of failure is less than 0.04? The advantage of having driven the piles to a penetration of 105 m is that additional information is available about their axial capacity from the pile driving resistance. Based on an analysis of the pile driving records, the capacity of the short piles is estimated to be 17 MN. This estimate is considered to be unbiased and the c.o.v. in the capacity is reduced from 0.3 to 0.1 due to this additional in situ data. Therefore, the input to the reliability calculations is updated as follows:

$$\text{capacity}_{\text{design, short}} = 17\,\text{MN}$$
$$\text{load}_{\text{design}} = 13\,\text{MN}$$
$$\text{FS}_{\text{design}} = 1.31$$
$$\text{load}_{\text{median}}/\text{load}_{\text{design}} = 1.5$$

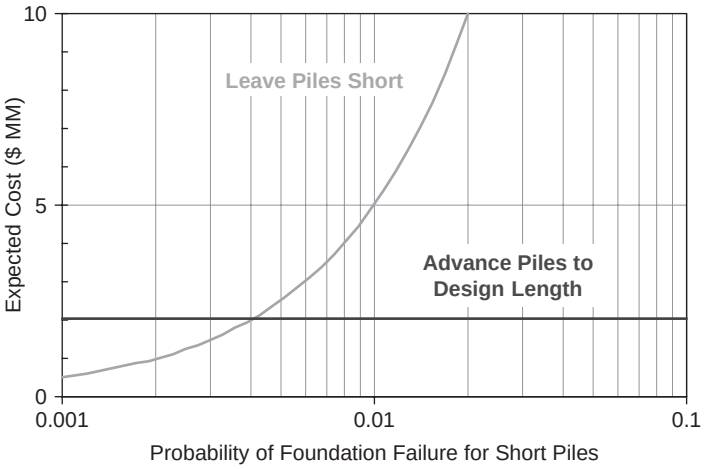


Figure 4 Comparison of expected costs for example with pile driving difficulty.

$\text{capacity}_{\text{median}}/\text{capacity}_{\text{design, short}} = 1.0$
$\text{FS}_{\text{median, short}} = 1.96$
$\delta_{\text{load}} = 0.3$
$\delta_{\text{capacity, short}} = 0.1$
$\delta_{\text{total, short}} = 0.32$

Hence, the probability of failure for an individual pile if it is left short of its design penetration is 0.02 from Eq. (2) or Fig. 3, which is less than the target value of 0.04.

This risk-based analysis supports a decision to leave the piles short. Note that the "design" factor of safety for the short piles is 1.31 and less than the design guidance of 1.5. However, there is less uncertainty in this estimate because it is based on in-situ, pile driving data. In addition, an important factor in making this decision is the cost associated with trying to advance the piles further (Fig. 2). If it will cost nothing, then it is clear that advancing the piles to their intended length is preferable. However, there is a need to balance costs associated with reducing the risk of failure. It is not possible to be conservative for the sake of being conservative because the cost of producing energy hangs in the balance.

2.2 *Floating production system*

A floating offshore structure was located in about 1,500 m of water and designed to gather oil production from a series of subsea well systems (Fig. 1). This example was first presented in Gilbert et al. (2008). The combination of load and resistance factors to be used for the design of the caisson foundation for this facility was called into question due to the unique nature of the facility: the loads on the foundation were dominated by the sustained buoyant load of the structure versus transient environmental loads, and

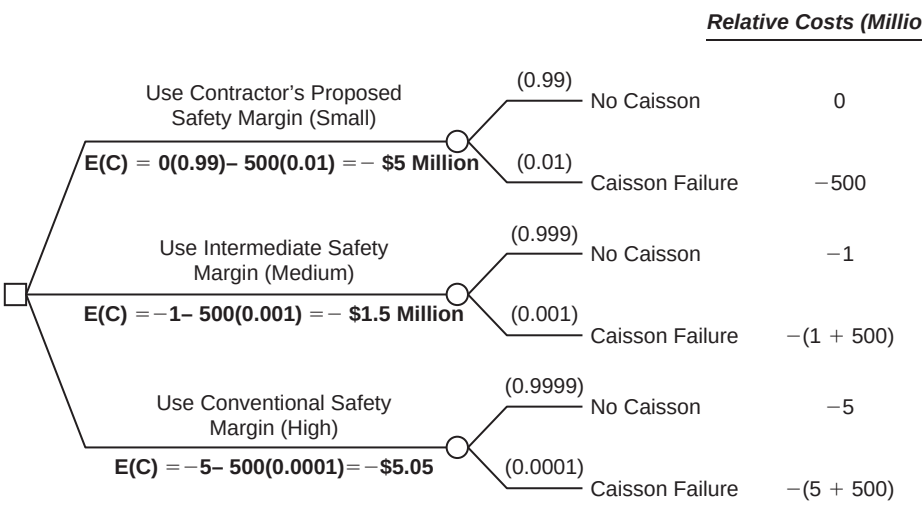


Figure 5 Decision tree for suction caisson design (from Gilbert et al. 2008).

the capacity of the foundation was dominated by its weight versus the shear resistance of the soil.

The design-build contractor proposed to use a relatively small safety margin, defined as the ratio of the load factor divided by the resistance factor, due to the relatively small uncertainty in both the load and the capacity. The owner wanted to consider this proposed safety margin together with a typical value for more conventional offshore structures, which would be higher, and an intermediate value.

The two considerations in this decision were the cost of the foundation, which was primarily affected by needing to use larger vessels for installation as the weight and size of the caisson increased, and the cost associated with a pull-out failure of the caisson if its capacity was not adequate. The decision tree in Fig. 5 shows the structure of decision alternatives and outcomes that were considered.

The probabilities and costs for caisson failure correspond to the event that the caisson will pull out of the mudline at some time during the 20-year design life for the facility. The cost of the caisson increases significantly if the highest safety margin is used because a more costly installation vessel will be required. In this example, the minimum expected cost is obtained for the intermediate safety margin (Fig. 5). Note that the contribution of the expected cost of failure to the total expected cost becomes relatively insignificant for small probabilities of failure.

In risk-based design decisions, there is also a consideration of achieving a minimum level of reliability that is considered to be "acceptable" by industry or society. In terms of the decision trees in Figs. 2 and 5, this consideration means adding an additional alternative of abandoning the facility if this minimum reliability cannot be achieved practically. Fig. 6 shows information for this minimum reliability level for offshore production facilities. Note that the contractor's proposed design approach in Fig. 5 would have achieved an "acceptable" level of reliability according to this industry

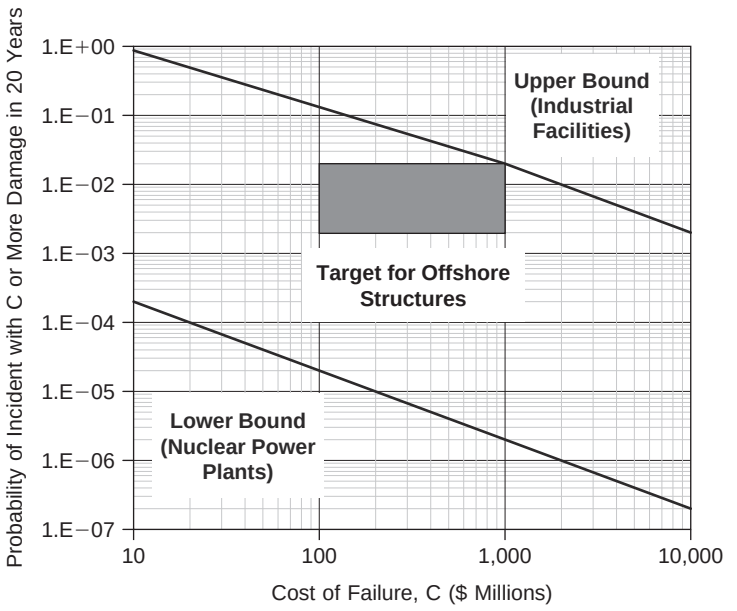


Figure 6 Acceptable risk criterion for evaluating offshore structures (from Gilbert et al. 2008).

standard. However, project-specific considerations of the risk of failure and the cost of reducing this risk lead to a higher level of reliability being preferred.

3 Is more information needed?

The value of additional information depends on how the information affects decisions that are made using the information. If a design will be same whether or not the information is obtained, then there is essentially no value to obtaining the information. A common occurrence of this question in practice is whether or not additional soil borings are needed to design the foundation for a production facility.

Consider a proposed jacket offshore structure (Fig. 1). A soil boring was drilled at the preliminary location of the structure. Subsequently, the location of the structure was moved about 500 m from the location of the soil boring. Do we need a soil boring drilled at the new location of the structure?

An example of an analysis to quantify the value of information in design for an oil production facility is shown on Fig. 7. Details for this example can be found in Gilbert et al. (2008). The geotechnical design properties derived from the boring data are not known before drilling the boring. There is a range of possible design properties that could be obtained from the boring, and, therefore, a range of possible pile designs that would be needed for a given set of geotechnical properties.

The major input to the decision tree on Fig. 7 is the magnitude of additional conservatism required from not having a site-specific boring. This conservatism can be

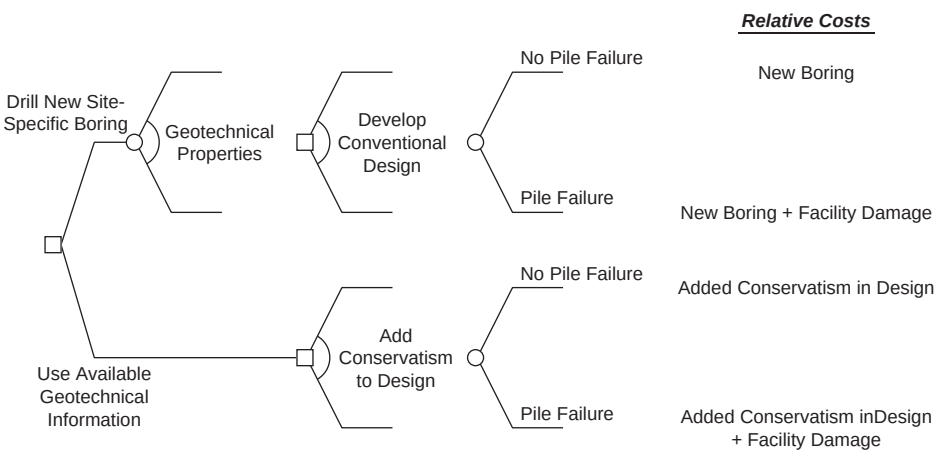


Figure 7 Decision tree for soil boring example (taken from Gilbert et al. 2008).

quantified using Eq. (2). The c.o.v. in the capacity is greater due to not having a site-specific soil boring:

$$\delta_{\text{capacity}} = \sqrt{\delta^2_{\text{design model}} + \delta^2_{\text{spatial variability}}} \tag{5}$$

where $\delta_{\text{design model}}$ is the conventional c.o.v. value of about 0.3 that accounts for uncertainty in the actual capacity compared to the design capacity when a site-specific soil boring is available, and $\delta_{\text{spatial variability}}$ is an additional c.o.v. value to account for not having a site-specific boring. The magnitude of $\delta_{\text{spatial variability}}$ can be quantified using available geologic and geotechnical data from the vicinity of the platform location (e.g., Gambino and Gilbert 1999).

If the same level of reliability is targeted whether or not a soil boring is available, then the increase in δ_{capacity} in Eq. (2) can be accounted for by increasing FS_{median} so that the same probability of failure is achieved. Practically, this compensation can be accounted for by increasing FS_{design} in Eq. (3) with a partial factor of safety:

$$FS_{\text{design}} = FS_{\text{no boring}} \times FS_{\text{design, boring}} \tag{6}$$

where $FS_{\text{design, boring}}$ is the conventional factor of safety used when a site-specific boring is drilled and $FS_{\text{no boring}}$ is a partial factor of safety that accounts for the additional uncertainty in not having a soil boring. This additional conservatism can also be expressed as a partial resistance factor in a load and resistance factor formulation.

The required value of additional conservatism is shown on Fig. 8 for this example field. The magnitude of $\delta_{\text{spatial variability}}$ depends on the distance away from the available soil boring (e.g., Gambino and Gilbert 1999, Gilbert et al. 1999 and Gilbert et al. 2008). At a distance of 500 m, $\delta_{\text{spatial variability}}$ is 0.07, meaning that the required partial factor of safety, $FS_{\text{no boring}}$, is 1.1 (Fig. 8).

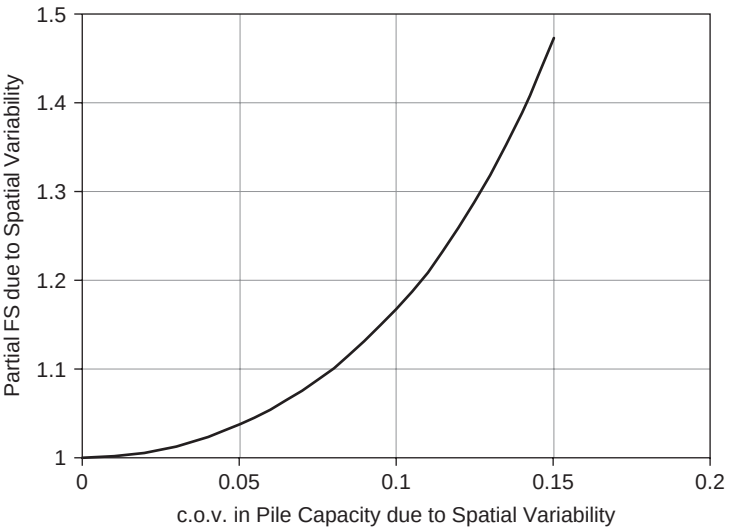


Figure 8 Required partial factor of safety with no site-specific soil boring versus magnitude of uncertainty due to spatial variability.

The value of drilling an additional soil boring can now be answered using the decision tree in Fig. 7. In this example, the cost of a new soil boring is $500,000; this cost includes the direct costs of drilling the boring and also the indirect costs of delaying installation until the boring is drilled. Increasing the factor of safety by ten percent corresponds approximately to increasing the length of the piles by about ten percent. The piles designed at the location of the existing boring were 100-m long; hence, the cost of not drilling an additional boring is the cost to increase each pile length by 10 m, which is $25,000 per pile. For four piles, the alternative of proceeding without a site-specific boring is going to cost $100,000 more than if the locations of the structure and the soil boring coincided. Since it will cost $500,000 to drill a new soil boring, the alternative of proceeding without the boring is preferred.

4 How can different types of information be incorporated into a decision?

A challenge in making a decision is to incorporate all available sources information, including measured data of varying qualities and types as well as subjective information. An example of environmental restoration at a large and old refinery (Fig. 9) provides an illustration of how different types of information can be utilized. Details for this example are provided in Gilbert (2002). This refinery is active, it is several hundred hectares in area, and it has been operating since the early 1900's. At present, there are numerous sources of contamination from past activities that will require corrective action over the next decade. Sources include contaminated soils, pools of nonaqueous phase liquids floating on the water table, and contaminated groundwater.

Figure 9 Aerial photograph of oil refinery (taken from Gilbert 2002).

One contaminant of concern was benzene in the surface soil near storage tanks and processing facilities. The types of information available to assess this contamination were (1) historical information about land use; (2) documented releases; and (3) measured benzene concentrations.

The approach taken to assess the extent of contamination was to first divide the facility into a set of small areas that were distinguished by land use, such as the secondary containment berm around each petroleum storage tank (Fig. 9). The next step was to estimate the fraction of soil volume within an area that had a concentration greater than the regulatory limit, 50 mg/kg. This fraction of "contaminated" soil was estimated using historical records for land use and product releases. A Beta probability density function (Fig. 10) was used to represent the distribution of possible fractions for contaminated soil as a function of the information available:

$$f_H(h) = \frac{\Gamma(q+r)}{\Gamma(q)\Gamma(r)} h^{q-1}(1-h)^{r-1} \tag{7}$$

where h is the fraction of contaminated soil, $\Gamma(\cdot)$ is the gamma function, and q and r are parameters that describe the shape of the distribution. The available information for a particular area was divided into one of four categories (Fig. 10). "Unlikely Contamination" was assigned to areas where no operations with benzene were ever conducted (stored or processed) based on historical land use; "Possible Contamination" was assigned to areas where operations with benzene were present historically, such as a storage tank; and "Likely Contamination" was assigned to areas where there were documented releases of benzene.

Once the initial or prior probability distribution for the fraction of "contaminated" soil was established based on the historical records, this distribution was updated with all available measurements of benzene concentrations in soil samples. Bayes' theorem

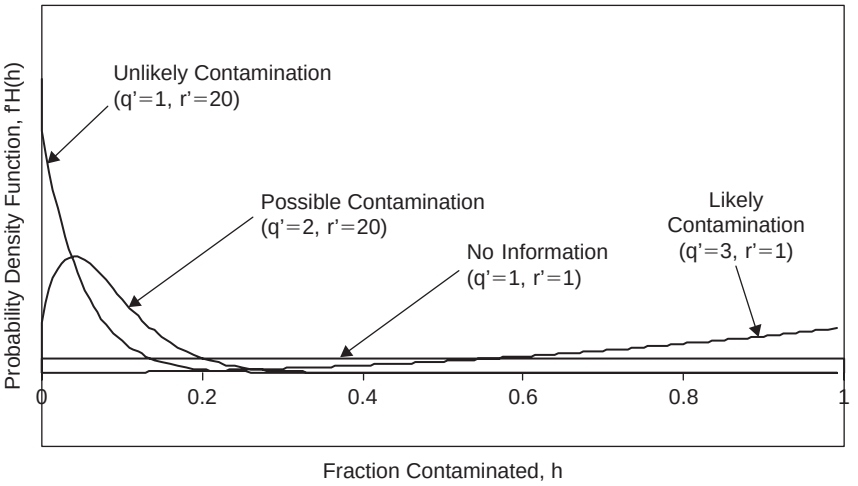


Figure 10 Prior distributions for fraction of contaminated soil based on historical information.

was used to update the prior distribution with the measured data as follows (Ang and Tang 2007):

$$P(H = h|\text{test results}) = \frac{P(\text{test results}|h)P(H = h)}{P(\text{test results})} \tag{8}$$

where $P(H = h|\text{test results})$ is the updated probability of interest for possible fractions of contaminated soil, $P(H = h)$ is the prior probability for possible fractions and obtained from the prior distribution (Fig. 10), $P(\text{test results}|h)$ is the probability of obtaining x contaminated results in n tests and given by a Binomial distribution assuming that the individual test results are statistically independent, and P(test results) is obtained from the Theorem of Total Probability by summing the product $P(\text{test results}|h)P(H = h)$ over all possible values for h. An advantage of using a Beta distribution for the prior probability is that an analytical solution is available for the updated probability distribution (e.g., Ang and Tang 2007). The updated probability is also a Beta distribution (Eq. 7) with the following parameters:

$$\begin{aligned} q'' &= q' + x \\ r'' &= r' + n - x \end{aligned} \tag{9}$$

where q', r' and q'', r'' are the parameters describing, respectively, the prior and updated Beta distributions. An example of the initial and updated Beta distribution is shown in Fig. 11 for an area where there was a storage tank but no historically documented spills. One of five soil samples taken from this area was contaminated with measured benzene concentrations greater 50 mg/kg.

The resulting probability distribution for the fraction of contaminated soil was put into more physically meaningful terms to support decision making. It was determined

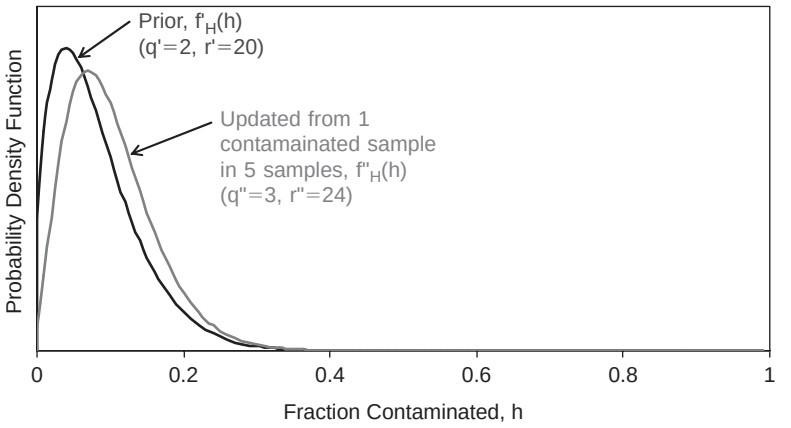


Figure 11 Prior and updated distributions for fraction of contaminated soil in area with storage tank.

that regulators would require 20 soil samples to be taken from an area with possible benzene contamination, and that they would conclude that no further action was required only if all 20 samples had concentrations less than 50 mg/kg. The implied criterion being used by the regulators is that they would not require corrective action if the fraction of contaminated soil in an area is less than 1 in 20, or 0.05 of the soil volume. However, there is a level of confidence implied in making this determination on the basis of 20 samples. If 20 samples are randomly selected from an area where no other information is available, then there is a 66-percent (not 100-percent) probability that the fraction of contaminated soil is actually less than 0.05 of the volume.

The updated probability distributions for the fraction of contaminated soil in each area could then be expressed in a manner to determine whether correction action would be necessary. The probability that the fraction of contaminated soil is less than 0.05 is shown in Fig. 12. If this probability is greater than 66 percent, then corrective action is not required (Fig. 12). An area could require corrective action if it is likely that there are high benzene concentrations in the soil. However, an area could also require corrective action if little is known about it, meaning that the distribution for the fraction will be wide and the probability that the fraction exceeds 0.05 will be relatively high. In many areas, there are few if any measurements of benzene concentration, and the subjective information based on historical site use is the only available information.

Additional sampling of soil is most valuable in areas where the information is most likely to change the decision. For example, consider an area where the probability that the fraction is less than 0.05 is 55%. Since this probability is smaller than 66%, this area would require corrective action if no additional information were obtained. However, since the probability is close to 66%, there is high likelihood that additional measurements could raise the probability to greater than 66% and consequently eliminate the cost for corrective action in this area.

An example of the value of information for two areas is shown on Fig. 13; the value of the information is quantified as a percentage of the corrective action (or remediation)

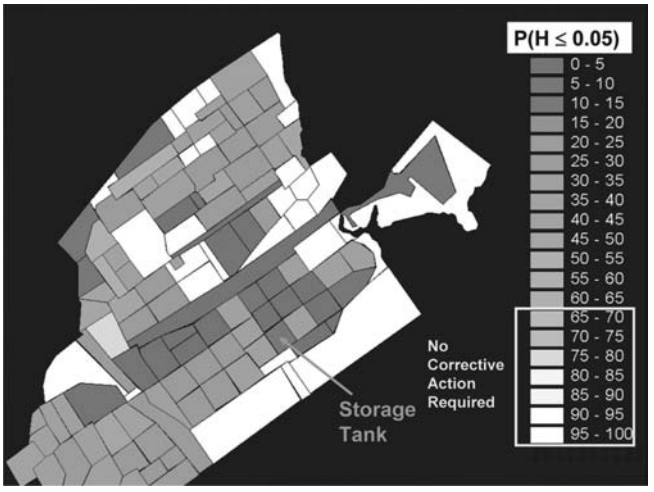


Figure 12 Updated probability that fraction of contaminated soil is less than 0.05 versus area of refinery based on soil sampled data combined with historical information (adapted from Gilbert 2002).

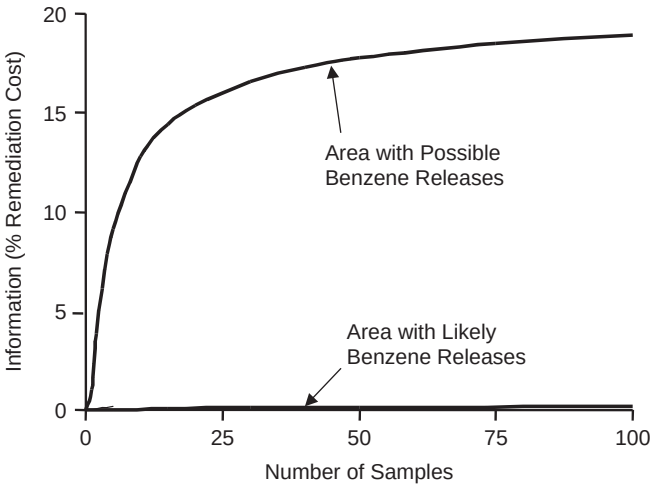


Figure 13 Value of information versus number of additional soil samples for two different areas (taken from Gilbert 2002).

cost. For an area with likely benzene releases, obtaining additional measurements is of limited value because it is very likely that these additional measurements will just confirm that corrective action is needed (Fig. 13). However, in an area with possible benzene releases, the value of the information is greater because it is likely that the cost of corrective action can be saved (Fig. 13). The value of the information can then be compared with the cost of obtaining additional data to decide how much (if any) additional data are warranted.

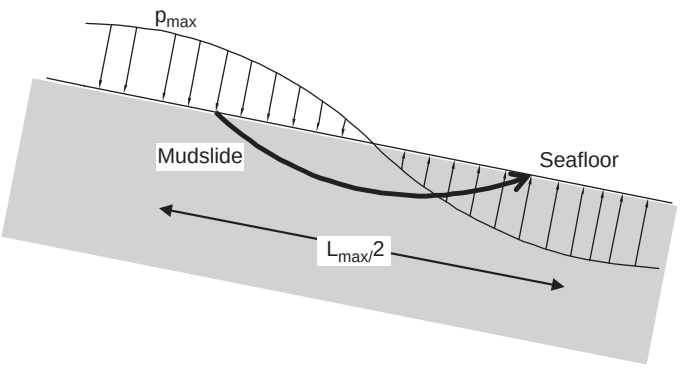


Figure 14 Schematic of wave-induced bottom pressures.

5 How can risk be assessed?

The risk of failure is an important component in decision making. In the basic decision framework (e.g., Fig. 2), the risk of failure is the expected loss or cost of a failure:

$$\text{Risk} = \text{Cost}_{\text{failure}} \times \text{Probability}_{\text{failure}} \tag{10}$$

More generally where there are multiple possible modes of failure, the risk is expressed as:

$$\text{Risk} = \sum_{\substack{\text{all i} \\ \text{failure} \\ \text{modes}}} \text{Cost}_{\text{mode i}} \times \text{Probability}_{\text{mode i}} \tag{11}$$

An example of assessing the risk for offshore oil pipelines illustrates the significance of both probability and cost. Details for this example are provided in Nodine et al. (2008).

The Gulf of Mexico has about 10,000 km of offshore pipelines that carry oil and gas from offshore production facilities to onshore refineries. Because there are numerous refineries located along the Louisiana coast near the Mississippi River, significant pipelines cross the Delta of the Mississippi River. In recent hurricanes in 2004 and 2005, a number of these pipelines failed due to wave-induced mudslides in the soft sediments that comprise the Delta. These failures shut in a substantial portion of offshore production in the Gulf of Mexico for months and even years. Therefore, assessing the risk of future disruptions is of interest in deciding how best to manage this risk.

The probability of a submarine mudslide depends on the characteristics of wave-induced pressures on the ocean bottom (the hazard) and the potential for failure of the seafloor under a given distribution of bottom pressure (the vulnerability). The hazard is characterized by a wave amplitude and a wave length for the distribution of bottom pressure corresponding to an instant in time at a given water depth (Fig. 14). The probability for a particular combination of wave height and amplitude, $P(p_{max}, L_{max})$,

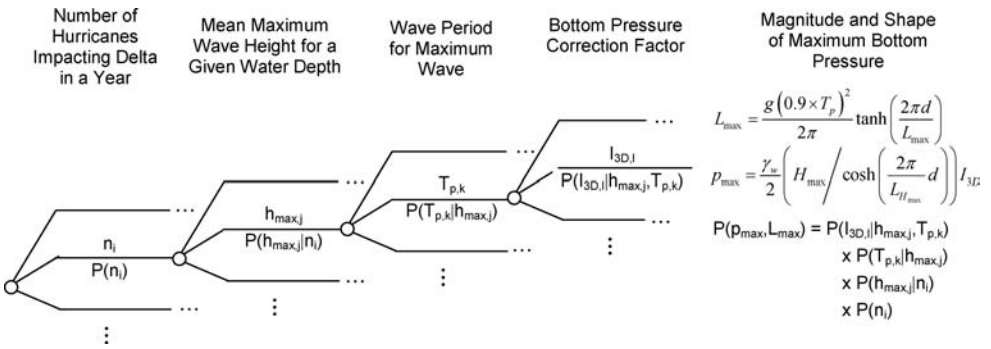


Figure 15 Event tree to assess the hazard for wave-induced bottom pressures.

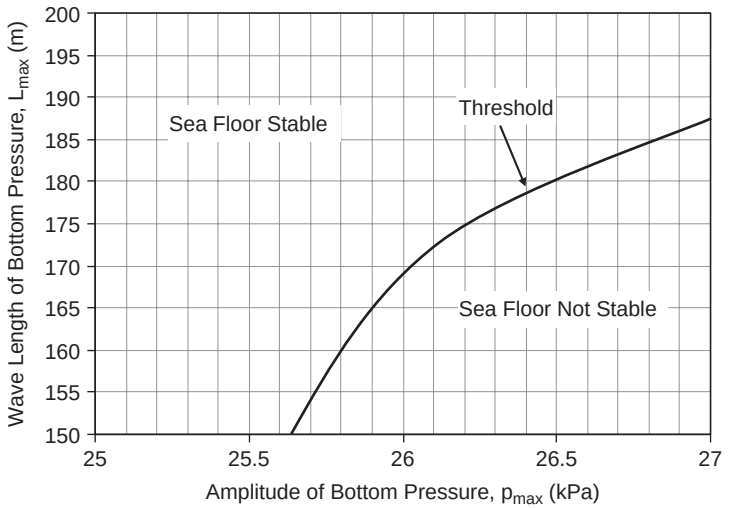


Figure 16 Vulnerability of mudslide versus wave amplitude and length for bottom pressure for a particular water depth and bottom slope.

is assessed using the event tree in Fig. 15. This event tree depicts the important factors and their inter-relationships in characterizing the hazard for wave-induced mudslides. The conditional probabilities in this event tree, such as the probability that the wave period is equal to a particular value given a maximum wave height or $P(T_{p,k}|h_{max,j})$, are assessed based on historical data and modeling for hurricanes.

The vulnerability at a particular location is expressed in terms of the combination of bottom-pressure wave amplitude and length that will cause failure for a given profile of soil strength versus depth below the sea floor (Fig. 16). The threshold of instability on Fig. 16 depends on the slope of the sea floor at the particular location. For a given profile of shear strength, the probability of a slope failure is obtained by summing up the probabilities for all possible values of p_{max} and L_{max} that are below the threshold in

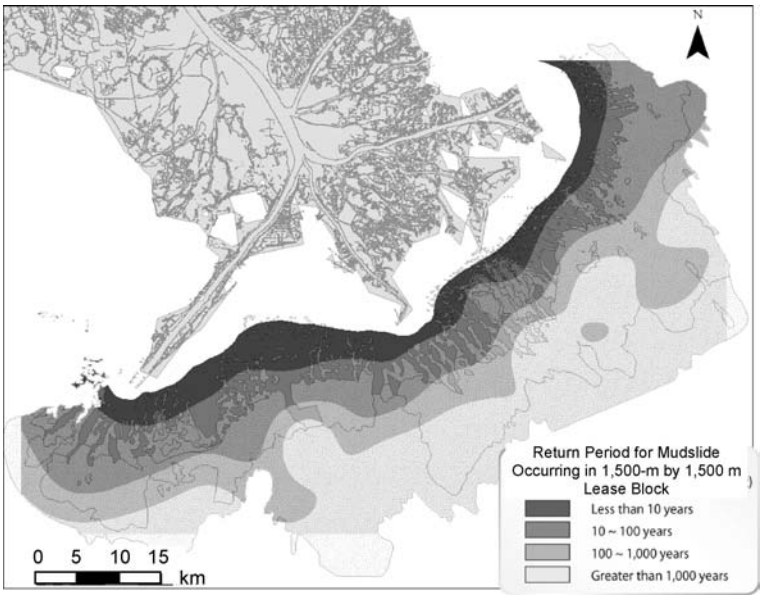


Figure 17 Return period for wave-induced mudslides in Mississippi Delta (from Nodine et al. 2008).

Fig. 16, where these probabilities depend on the water depth at the particular location. Finally, since the properties of the sea floor vary significantly across the Delta and are not generally known at a particular location, the spatial variability in sea floor properties is considered by accounting for a range of possible profiles of soil shear strength versus depth at each location. The probability of failure for each possible profile of soil shear strength is multiplied by its probability of occurrence at a random location in the Delta, and the product is summed for all possible profiles of soil strength through the Theorem of Total Probability.

An example set of results is shown in Fig. 17. In this figure, the return period is shown for a mudslide occurring within an area that is about 1,500 m square (approximately the size of a lease block). There are two notable features of these results. First, the probability of a mudslide depends on the location, specifically the water depth and the slope of the sea floor. Second, the return periods for mudslides at some locations in the Delta are less than 10 years, meaning that the risk of a pipeline failure in these locations can be high.

The cost of pipeline failures depends on how many and which pipelines are affected. One of the lessons from recent hurricanes in the Gulf of Mexico is the potential cost associated with a disruption to the oil distribution system. If multiple pipelines fail, then there may be no way to deliver the product from the production facilities to the refineries. The economic consequences of the pipeline failures in 2004 and 2005 are still being realized today in 2008. The recent increase in the world-wide price of crude oil was triggered and amplified to some extent by the loss of production caused by these hurricanes.

Assessing the risk of failure requires a careful consideration of the consequences or costs of component failures on large, inter-dependent systems for energy production, processing and distribution. Since these systems can have a variety of owners, operators and jurisdictional authorities, assessing the risk of failure may require a much broader perspective than that of an individual entity. Also, these systems change with time. Finally, the associated uncertainty in estimating these costs can be large and needs to be considered in assessing the risk. The form of Eq. (11) and event trees such as Fig. 15 allow for a variety of different scenarios to be considered in how the failure of pipelines might lead to different consequences depending on how many pipelines fail, which pipelines fail and when the failures occur.

6 How does risk impact public policy?

Risk, whether real or perceived, has a significant impact on public policy. As new technology is developed for producing conventional and alternative sources of energy, public policy will play an important role in either promoting or stifling these developments. Hence, a rational assessment and analysis of the risk associated with new technology is needed to maintain a reasonable balance of risks, costs and benefits.

As an example of how risk can impact public policy, the United States government instituted a policy disallowing the use of shuttle tankers to transport oil from offshore production facilities in the wake of the oil spill from the Exxon Valdez tanker. However, as we have moved into deeper water and further offshore in producing oil, the cost of extending and maintaining the network of offshore pipelines has increased substantially. In addition, as experienced in recent hurricanes and demonstrated in Fig. 17, there is a risk of failures for pipelines due to submarine mudslides (along with other causes). Therefore an analysis was conducted for the government to compare the risks of oil spills associated with the conventional technology for oil transportation, pipelines, and an alternative technology, shuttle tankers (Gilbert et al. 2001).

The process used for this risk analysis was as follows. First, a metric was chosen to measure risk: the total volume of oil spilled into the water for that particular oil transportation technology over a 20-year lifetime for a production facility in deepwater. This metric was chosen based on the following rationale:

- Spill volume provides relevant and useful input to the government in their decision-making process;
- Spill volume is tractable and quantifiable; and
- Spill volume is currently tracked and recorded so that (i) available data could be used to support the results of the risk analysis and (ii) future data can be used to validate and calibrate the results.

Second, a set of representative study systems was defined for comparison purposes (Fig. 18): three conventional deepwater production and transportation systems, a floating Tension Leg Platform (TLP) with oil pipelines, a floating Spar with oil pipelines, and a fixed jacket in shallow water that serves as a host for subsea well systems in deepwater and as a hub for oil pipelines from deepwater; and one system representing new technology, a Floating Production, Storage and Offloading (FPSO) system where the oil is stored in field offshore and transported to shore via shuttle tankers.

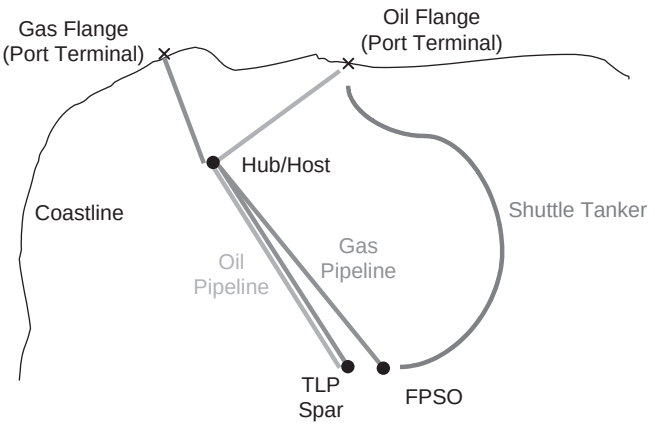


Figure 18 Study oil transportation systems for comparative spill risk study (taken from Gilbert et al. 2001).

Third, all of the available data on spill frequencies and sizes were compiled and analyzed for similar systems in the Gulf of Mexico and around the world. Fourth, a series of workshops was held with representatives from oil companies, consultants, manufacturers, contractors, classification societies, and regulatory agencies to sort through and extrapolate from the data to predict how the study systems would perform. Finally, the input was used to assess and analyze the risk of oil spills from each study system.

The results of this risk analysis are shown in Figs. 19 and 20. The annual frequencies for spills of different sized from each study system are shown in Fig. 19; expected values and 90-percent confidence bounds are shown for each frequency. The estimated frequency of spills tends to decrease as the spill size increases. Also, the relative magnitude of uncertainty in the estimated frequency increases as the spill size increases. This relative increase in uncertainty occurs because large spills are rare events, so there are few occurrences available from which to estimate frequencies. The new technology is estimated to have less frequent but larger spills than the existing technology (Fig. 19).

Estimates for the average total volume of oil spilled over a 20-year lifetime are shown in Fig. 20 for each study system. Given the uncertainty in the estimated performance, it is not possible to distinguish the oil spill risk between these various systems.

The information from this risk analysis was used together with other analyses by the government to re-visit the policy disallowing the use of shuttle tankers. In 2001, a Record of Decision was issued to modify this policy and allow for the use of shuttle tankers. There are now numerous production systems in development for the Gulf of Mexico that will rely on shuttle tankers. Furthermore, the risk associated with mudslides (Fig. 17) is spurring the development of a floating storage facility located outside of the Delta so that oil could still be transported through the Delta with shuttle tankers in the event of future pipeline failures.

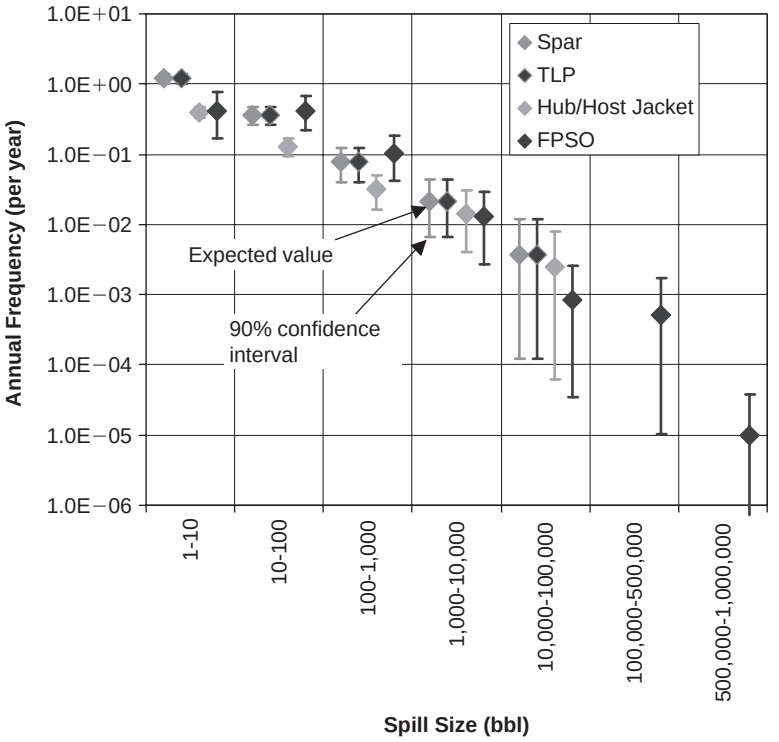


Figure 19 Results of spill frequency for study oil transportation systems (taken from Gilbert et al. 2001).

7 How best to communicate risk?

Communication of risk is important both to elicit input information and to communicate results to decision makers and stakeholders. People generally are not well versed in the meaning of statistics and probability, so numerical measures of uncertainty and risk are not effective in communicating risk. In addition, once information is presented, such as the interpolated cross-section in Fig. 21, it tends to be interpreted as reality and certainty no matter how many qualifications are provided to account for uncertainty in the information. It is imperative that engineers work with specialists in communication, sociology and psychology to achieve effective communication of risk and risk-based decisions (e.g., Mileti 2007).

An example means for effective communication of information is a graphical technique called multiples (Tufte 1990). Multiples are small-scale images positioned within the eye span on a single page or screen showing the range of possible interpretations of the information. Uncertainty multiples use this technique to convey uncertainty (Gilbert et al. 2006). To illustrate, uncertainty multiples are shown in Fig. 22 to convey uncertainty in the interpolated conditions between the two borings in Fig. 21. These multiples show that there is uncertainty in the presence and thickness of a buried alluvial channel at the location of a proposed offshore structure. The possibility of

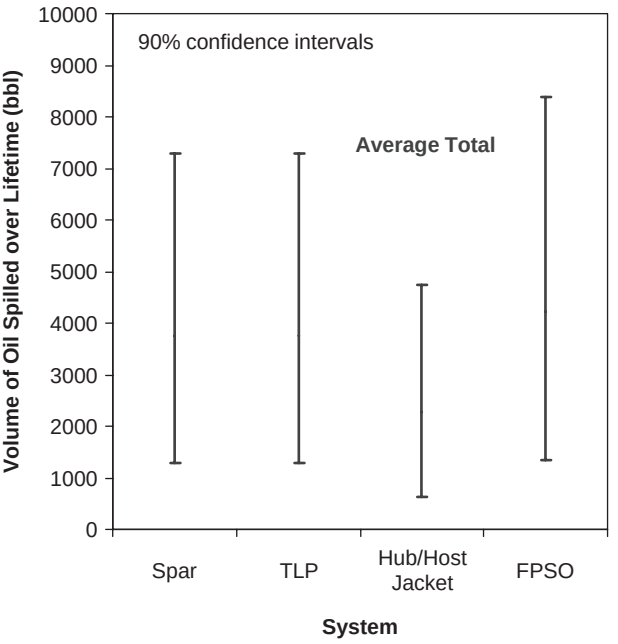


Figure 20 Results of predicted average volume of oil spilled for study oil transportation systems (taken from Gilbert et al. 2001).

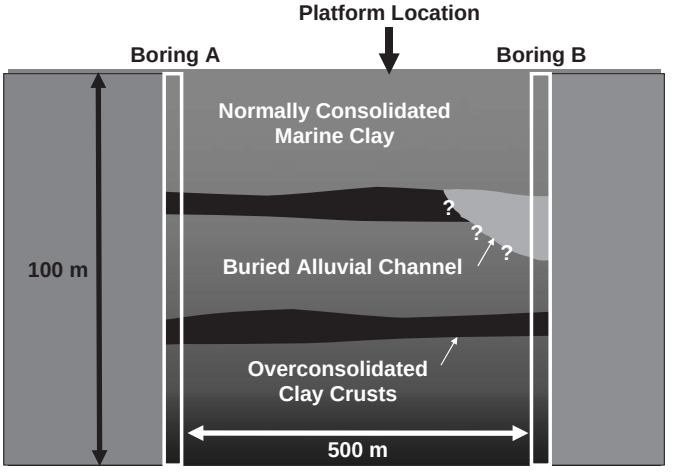


Figure 21 Example of interpolating between soil borings (adapted from Gilbert et al. 2006).

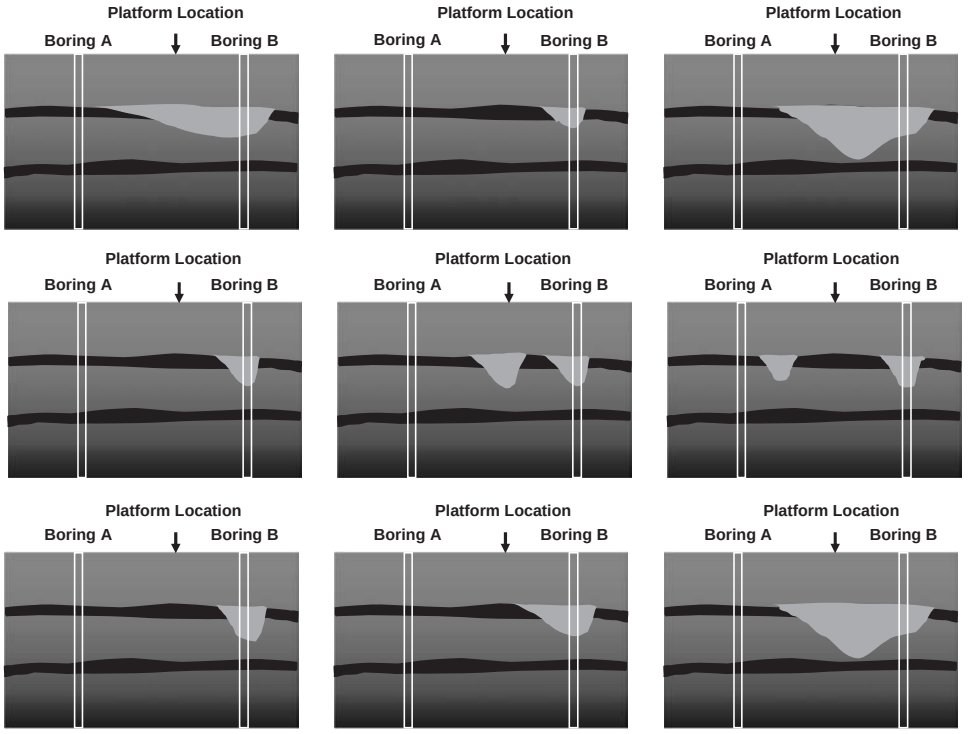


Figure 22 Example of uncertainty multiples to convey uncertainty in conditions between soil borings (taken from Gilbert et al. 2006).

encountering a buried channel is an important consideration for the platform designer and installer because it could slow or prevent pile driving.

Figure 22 shows a wealth of information about the uncertainty in encountering a buried channel at the platform site. Each small-scale image on Fig. 22 represents an equally likely scenario in the range of possibilities. In this way, the probability of a scenario is conveyed directly to the viewer. For example, there is a 4/9 probability that a buried channel will be encountered at the platform site. It may be the same channel as that encountered at Boring B or a different channel. Also, there is a small chance (1/9) that the thickness of the channel at the location of Boring B is actually smaller than was measured due to sand caving into the hole during drilling (middle image in upper row of Fig. 22). Finally, if a channel is encountered its thickness is uncertain and ranges between 10 and 30 m. In summary, Fig. 22 shows more information to the decision makers in a single glance than could possibly be conveyed in statistical parameters or in paragraphs of text.

8 Summary

A risk-based decision framework is valuable in developing and maintaining the infrastructure for producing, processing and distributing energy. This framework helps to

provide a context so that costs and benefits can be balanced in managing risks. There are several important principles that can be derived from this framework:

1. Excessive conservatism is not practical and the benefits of reducing risk need to be balanced against their costs. The optimal degree of conservatism is best considered within a project-specific context.
2. The value of obtaining additional information to support decision making depends on the potential for the information to change decisions and needs to be balanced against the cost of obtaining the information.
3. All available information can and should be incorporated into a risk-based decision.
4. Assessment of risk requires a careful consideration of both the consequences of a loss and their respective probabilities.
5. Rational and objective assessments of risk are needed in formulating public policy.
6. Effective communication of risk and risk-based decisions is essential.

The real benefit of a risk-based approach is in providing a logical, consistent and comprehensive platform to discuss and evaluate important decisions. It is the process, and not the quantitative results, that should be embraced and emphasized. The most effective risk-based approach involves engineers reaching out to and working with the public as well as a variety of other professionals, ranging from policy makers to sociologists and economists.

Acknowledgment

I wish to acknowledge the following organizations for supporting the research and consulting upon which this paper is based: American Petroleum Institute, American Society of Civil Engineers, BP, CALFED, ExxonMobil, National Science Foundation, Offshore Technology Research Center, Shell, State of Texas Advanced Research and Technology Programs, Unocal, US Army Corps of Engineers, and US Minerals Management Service. I would also like to acknowledge the support of my colleagues and students at The University of Texas at Austin, as well as the Hudson Matlock Professorship Fund. The views and opinions expressed herein are my own and do not reflect any of the organizations, parties or individuals with whom I have worked.

References

Ang, A.A-S. & Tang, W.H., Probability Concepts in Engineering, 2nd Edition, J. Wiley & Sons, Inc., New York, 2007

Ang, A.A-S. & Tang, W.H., Probability Concepts in Engineering Planning and Design: Volume II – Decision, Risk and Reliability, J. Wiley & Sons, Inc., New York, 1984

Benjamin, J.R. & Cornell, C.A., Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill, Inc., New York, 1970

Gambino, S.J. and Gilbert, R.B., Modeling Spatial Variability in Pile Capacity for Reliability-Based Design, Analysis, Design, Construction and Testing of Deep Foundations, Roesset Ed., ASCE Geotechnical Special Publication No. 88, 135–149. 1999

Gilbert, R.B., Gambino, S.J. and Dupin, R.M., Reliability-Based Approach for Foundation Design without Site-Specific Soil Borings, *Proceedings*, OTC, Houston, Texas, OTC 10927, 631–640, 1999

Gilbert, R.B., Questions for the Future of Risk Assessment in Environmental Geotechnics, *Proceedings*, 4th Intl. Congress on Environmental Geotechnics, Rio de Janeiro, Brazil, Vol. 2, 1007–1012, 2002

Gilbert, R.B., Najjar, S.S., Choi, Y.J. and Gambino, S.J., Practical Application of Reliability-Based Design in Decision Making, in *Reliability-Based Design in Geotechnical Engineering: Computations and Applications*, Phoon Ed., Taylor & Francis Books Ltd., 2008

Gilbert, R.B., Tonon, F., Freire, J., Silva, C.T. and Maidment, D.R., Visualizing Uncertainty with Uncertainty Multiples. *Proceedings, GeoCongress 2006*, ASCE, Reston, Virginia, 2006

Gilbert, R.B., Ward, E.G. and Wolford, A.J., A Comparative Risk Analysis of FPSO's with other Deepwater Production Systems in the Gulf of Mexico, *Proceedings*, OTC, Houston, Texas, OTC 13173, 2001

Mileti, D.S., Public Hazards Communication and Education: The State of the Art, www.incident.com/access/images/6/68, 2007

Nodine, M.C., Wright, S.G., Gilbert, R.B., Ward, E.G., Cheon, J.Y., Mudslides During Hurricane Ivan and an Assessment of the Potential for Future Mudslides in the Gulf of Mexico," Final Project Report, OTRC, Prepared for Minerals Management Service, Herndon, Virginia, 2008.

Tang, W.H. and Gilbert, R.B., Case Study of Offshore Pile System Reliability, *Proceedings*, OTC, Houston, Texas, 677–686, 1993

Tufte, E.R., Envisioning Information, Graphics Press, Cheshire, Connecticut, 1990

Chapter 8

The role of risk assessment in performance-based engineering

Bruce R. Ellingwood

School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT: The design of civil infrastructure facilities for conditions imposed by their service requirements and by natural environmental events is guided by codes, standards and other regulatory documents. While constructed facilities so designed usually possess a degree of integrity that is also available to withstand challenges from unforeseen events, events not explicitly considered may lead to unacceptable economic losses or precipitate a catastrophic collapse. Structural engineers now are seeking improvements to building practices to achieve performance beyond what is provided by current prescriptive code minimums, to enhance facility robustness, and to mitigate unacceptable economic damages from low-probability, high-consequence hazards. The new paradigm of performance-based engineering (PBE) is evolving to enable structural design to better meet heightened public expectations and to achieve more reliable prediction and control of infrastructure performance. Uncertainties in structural loads, system strengths and stiffnesses, and other factors in the building process give rise to risk, which is managed through building codes, standards and other regulatory documents. This paper explores some of the issues raised by the use of structural reliability and risk assessment methodologies as decision support tools for dealing with uncertainties rationally in PBE.

1 Introduction

Structural codes and standards provide the foundation of good engineering practice for the design and construction of civil infrastructure. Recent advances in the science underlying structural design have made it possible to predict the behavior of complex structural systems quite accurately, provided that the loads and strengths of construction materials are known. Despite these advances, structural loads, material strengths and numerous other sources of uncertainty, both aleatoric (inherently random) and epistemic (knowledge-based) in nature, remain in the building process. These uncertainties are addressed with various safety factors in codes and standards. The safety level inherent in many codes represents a value judgment by the code-writers based on past experience. This judgmental approach to safety assurance served the profession well in an era when design and construction technologies were stable or evolved slowly and there was opportunity to learn from experience. In recent years, however, society and rapid evolutions in building technology have allowed less opportunity for learning by trial and error.

Structural design codes in most of the world remain prescriptive in nature, providing detailed criteria that define what is believed essential for adequate safety in the completed project. However, prescriptive criteria create the illusion that meeting the code

minimums will result in a satisfactory building or other structure. This is often not the case. The performance of many structures designed by traditional prescriptive criteria during extreme natural phenomena hazards has been inadequate (Ellingwood, 1998), the severe economic losses from natural disasters such as Hurricane Andrew and the Northridge Earthquake being two cases in point. At the same time, social conditions have evolved worldwide and public awareness and expectations regarding infrastructure performance and safety have increased markedly during the past thirty years as a result of media coverage of natural and man-made disasters. It has become apparent that while safety is an absolutely essential ingredient of acceptable performance, the prescriptive nature of traditional codes and their focus on life safety have made them less than effective in stemming the severe functional disruptions that such events may cause in buildings and other structures. Moreover, while the normal structural design process usually has resulted in a constructed facility with a degree of integrity that is available to withstand challenges from unforeseen events, there is increasing concern about the prospect of severe damage or catastrophic collapses brought about by extreme events that are outside the traditional design envelope.

Structural engineers now are seeking improvements to engineering practices to achieve building performance beyond what is provided by current prescriptive code minimums, enhance facility robustness and mitigate unacceptable economic damages from low-probability, high-consequence hazards. The recent interest in the structural engineering community in performance-based engineering (PBE) stems from a desire to tie the structural design process more closely to heightened client, owner and occupant expectations of infrastructure performance by achieving more reliable prediction and control of infrastructure performance (Hamburger, et al, 2003). In the past decade, the paradigm of PBE has achieved significant momentum in the United States, Canada, Western Europe, New Zealand and Japan, especially in the areas of earthquake engineering and fire protection engineering where there are strong economic motivations to adopt alternatives to the traditional prescriptive design approaches.

For performance-based engineering to achieve its full potential, however, the performance metrics and design criteria for different occupancy classifications and categories of construction must properly reflect the inherent and modeling uncertainties that govern structural performance. Recent advances in structural reliability and risk assessment applications now make performance classification on the basis of acceptable risk possible. Structural reliability theory has provided a quantitative link between the practice of structural engineering and its social consequences in the first-generation of probability-based limit states design (PBLSD) codes (Ellingwood and Galambos, 1982; Ellingwood, 1994), such as *ASCE Standard 7* (ASCE, 2005a). It promises to become even more important as a decision support tool in PBE, which must be supported by quantitative reliability and risk assessment tools. This paper summarizes some of the advances in probability-based code development that have been made possible by structural reliability methods and explores issues that must be addressed in the development and implementation of PBE.

2 Performance-based engineering

All modern performance-based code proposals have several basic features in common: a set of explicitly stated requirements or functional objectives arranged in a hierarchical

order, beginning with public safety and extending to include tolerable damages, economic loss, or loss of function; specific criteria needed to meet those requirements; evaluation methods by which satisfaction of the criteria can be measured; and extensive commentaries. Their aim is to couple the structural design requirements to performance expectations, to ensure that hazards are treated consistently and that design conservatism is appropriate to required function, and to expand the current focus of codes on safety to include the financial losses associated with failure to perform according to expectations.[1] Life safety remains the central objective, of course, but the degree to which other performance objectives are achieved may depend on the nature of the intended use of the facility, the degree of risk aversion of the owner, and the possible return on investment in design and construction in terms of additional performance during the service life of the facility. The developers of a PBE-based standard envision that the design team will actively discuss these various performance issues as part of the design development process.

The differentiation in levels of performance for different constructed facilities where life safety or economic consequences differ is one feature that distinguishes performance-based engineering from current codes of practice. It has been common in codes such as *ASCE Standard 7* (ASCE, 2005a) to require essential facilities to be designed for a higher nominal load. One would expect, e.g, that hospitals should be designed to a higher level of performance than one-story commercial establishments, and since the early 1970's, ordinary buildings have been designed for nominal (characteristic) wind loads that are calculated from a wind speed with return periods of 50 years, while the load used to design hospitals has been based on a wind speed with a return period of 100 years. Beyond this simple requirement, however, no quantitative relation has been established between structural design criteria and risk.[2] Most proposals for performance-based engineering take this relation to the next stage, through a matrix, such as that in Figure 1, which relates frequency of hazard and consequences of failure. It should be noted that current structural design criteria focus on the life *safety under rare events* element in Figure 1. The formulation remains semi-probabilistic; it (usually) models the hazard probabilistically (a rare event is one with a return period on the order of 1,000 years or, equivalently, an annual exceedence probability of 0.001), but does not admit the possibility that life safety may be endangered from events of lesser or greater intensity. Nor does it address the uncertainty in the capacity of the structure to withstand hazards of different intensity.

Structural reliability and risk assessment, supported by advances in computation and simulation, have advanced to the point where design for explicit levels of facility performance (e.g., life safety) can be related to a spectrum of hazards and risk-informed criteria can be developed to populate the boxes in Figure 1. The necessary ingredients in

[1] The concept of performance-based engineering would be familiar to structural engineers working in the aerospace, automotive and marine industries, where engineered products are mass-produced, demands vs capacities are relatively predictable, supporting data are available from component testing, extensive computational analysis is feasible, and the technology is more easily controlled through quality control/assurance. In such an environment, prescriptive structural design codes have a minimal role.

[2] In earthquake engineering (ASCE, 2005a), the structural systems permitted for specific occupancies and their detailing criteria are related to so-called *Seismic Design Categories*; however, the relation is qualitative in nature, and is not based on risk analysis.

		Consequence			
		Continued Service	Impaired Function	Life Safety	Incipient Collapse
Event magnitude	Small	All			
	Infrequent	III	II		
	Rare	IV	III	II	
	Extreme	NA	IV	III	II

Figure 1 Performance objectives and event probabilities for ASCE Standard 7 building categories (II – Normal occupancy; III – High occupancy; IV – Essential occupancy).

this assessment are probabilistic models of the spectrum of hazards and of the capacity of structural systems to withstand the structural actions resulting from those hazards. Classical system reliability formulations (e.g., Melchers, 1999) based on evaluating probabilities of unions of minimal cutsets have been largely superseded in the past decade by stochastic finite element methods that involve nonlinear dynamic response analysis and large-scale simulation.

3 Reliability bases for PBE

As noted above, differentiated levels of performance for various occupancy categories are characteristic of all proposals for performance-based design. These levels cannot be achieved consistently without a rational consideration of the role played by uncertainties due to occupancy or service environment, accidental conditions, and structural strength and stiffness. Structural reliability theory provides the framework for the rational evaluation of such uncertainties. The development of reliability-based performance criteria requires: identification of the various hazards that might challenge the system, probabilistic models that describe the relative likelihood of their occurrence; models that relate component behavior to system behavior (often, nowadays, through complex finite element analysis); and development of point or interval estimates of limit state (or damage state) probability that can be used for engineering decision analysis.

3.1 *Fundamental reliability analysis tools*

Each hazard is described probabilistically by a *hazard curve*, $H(x) = Prob[Q > x]$, which defines the probability that the random variable, Q, describing the demand

(force, velocity, spectral acceleration, depending on the nature of the analysis) exceeds a value, x. The capacity of a structural system to respond to and withstand challenges to its function or integrity is described by random variable, R, which must be dimensionally consistent with Q. The safety margin, M, then is defined as $M = R - Q$, and failure occurs when $M < 0$. In a fully coupled risk analysis, the limit state probability is obtained by,

$$P[LS] = \int_0^\infty H(x) dF_R(x) \tag{1}$$

in which $F_R(x) =$ cumulative distribution function (CDF) of R. For risk-informed decision-making in the presence of uncertainties, this limit state probability becomes the metric[3] for measuring the degree to which a specific performance objective is met. Frequently, a *damage state (DS)* can be associated with a limit state; for example, in *ASCE Standard 41-06* (ASCE, 2006b), the states of continued (or impaired) occupancy, life safety and incipient collapse (or collapse prevention) are associated with maximum inter-story drifts.[4]

It might be observed that $F_R(x)$ is the conditional limit state probability that $R \leq x$, conditioned on $Q = x$. This conditional probability is denoted a *fragility*. Risk analyses of structural systems (performed by first-order reliability analysis or, more recently, by Monte Carlo simulation coupled to nonlinear finite element analysis) have indicated that the fragility often can be described by a lognormal distribution (Singhal and Kiremidjian, 1996; Shinozuka, et al, 2000; Wen and Ellingwood, 2005):

$$P[LS|Q = x] = F_R(x) = \Phi[\ln(x/m_R)/\beta_R] \tag{2}$$

in which $m_R =$ median (system) capacity and $\beta_R =$ logarithmic standard deviation, which is a measure of uncertainty in capacity that is closely related to the coefficient of variation in R. A properly constructed fragility model of a structural system provides a measure of its likely performance under specified extreme events. Figure 2 is a schematic of a steel braced frame in an existing six-story building in Memphis, TN (Ellingwood, et al, 2007). Figure 3 illustrates the seismic fragilities for this frame for three damage states – Immediate Occupancy, Structural Damage, and Incipient Collapse, associated with maximum inter-story drifts of 0.01, 0.02 and 0.04, respectively – determined by simulation and non-linear finite element time history analysis, and plotted as a function of spectral acceleration at the fundamental period of the building (1.04s). One might assert from this analysis that given the occurrence of an earthquake with spectral acceleration of 0.3g, the probabilities of continued occupancy, impaired occupancy, severe structural damage and collapse are, respectively, 0.02, 0.68, 0.29 and 0.01. However, these probabilities are a less informative measure

[3] Other metrics for risk-informed decision-making include (annual) expected economic loss or expected mortality. However the limit state probability is the starting point for determining such losses, and its use avoids some of the social and political issues involved with the use of other metrics.

[4] The damage state probability is often more useful in communicating risk to a non-technically trained stakeholder or insurance underwriter than the limit state probability, which requires some knowledge of structural engineering concepts. In this paper, however, no distinction is made between them.

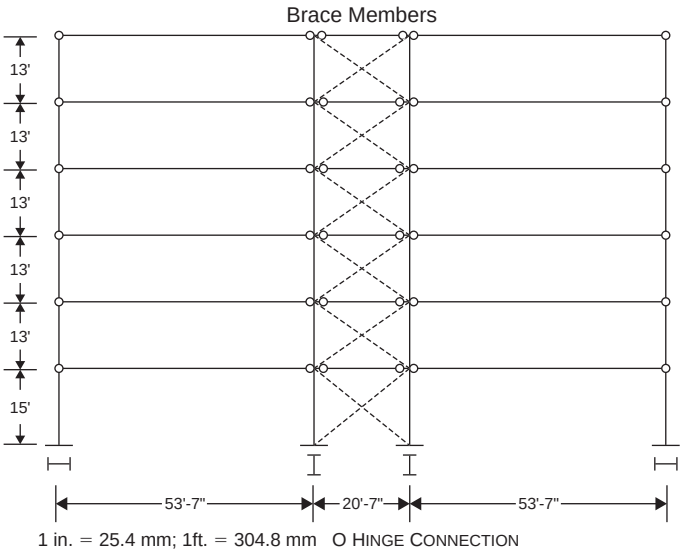


Figure 2 Schematic of six-story steel braced frame designed in 1989 in Memphis, TN.

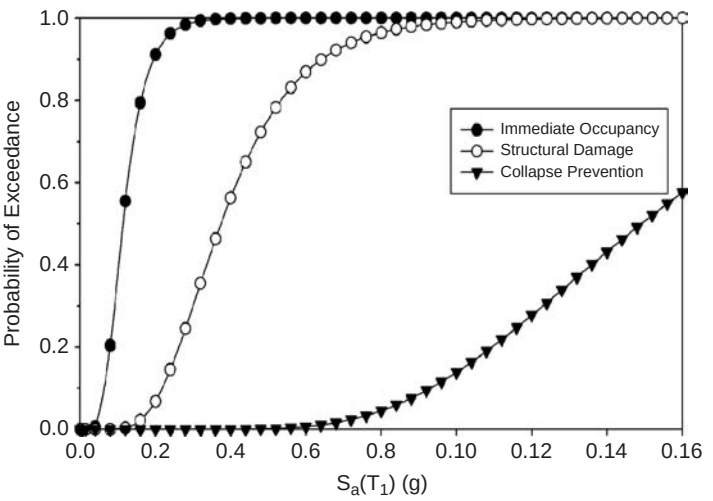


Figure 3 Seismic fragilities for six-story braced frame illustrated in Figure 2.

of risk than the limit state probability, $P[LS]$, in Eqn. 1, which covers the complete spectrum of hazards.

Appropriate probability models for natural or man-made hazards often are one of the distributions of extreme largest values, making Eqn 1 difficult to evaluate analytically. However, simplifications may be possible in some instances. For example, over the range of return periods of interest in structural engineering (100–10,000 years),

there often is a (nearly) logarithmic-linear relation between demand, x, and annual probability that specific levels are exceeded (Cornell, et al, 2002):

$$\ln H(x) \approx \ln k_0 - k \ln x \tag{3}$$

in which k_o, k = parameters. The parameter k measures the variability in hazard: $k = \ln(T_2/T_1)/\ln(x_2/x_1)$ for demands x_i and mean return periods T_i. If, e.g., doubling the demand increases the return period by a factor of 5, 10, and 100, respectively, then $\alpha = 2.32$, 3.32, and 6.64, respectively. If $F_R(x)$ and $H(x)$ are described by Eqns 2 and 3, respectively, then,

$$P[LS] = H(m_R)exp[(k\beta_R)^2/2] \tag{4}$$

in which $H(m_R)$ = hazard, evaluated at the median fragility, and the exponential term is a correction term to account for the slope of the hazard curve and variability in capacity. Equation 4 can be used to evaluate the risk implications for various performance-based code proposals based on different hazard levels, limit or damage states, and building performance categories.

3.2 *Uniform risk vs uniform hazard design*

Current design for extreme environmental hazards often is based on the notion of a "uniform hazard" or an event with geographically uniform probability of exceedence, provided in the form of a hazard map.[5] Suppose that the design resistance, R_d, is stipulated at the 5%ile of the fragility (cf Figure 3). The lognormal model of fragility then implies that,

$$R_d = m_R\, exp[-1.645\beta_R] \tag{5}$$

Solving Eq (5) for m_R and substituting in Eq (4), one obtains,

$$P[LS] = H(R_d) \exp[(k\beta_R)^2/2 - 1.645k\beta_R] \tag{6}$$

In words, Eq (6) says that the limit state probability equals the hazard exceedence probability evaluated at the design resistance, multiplied by a correction factor that is a function of the slope of the hazard curve (factor k) and the uncertainty in system capacity (factor β_R). This equation can be used to develop a *uniform risk* criterion for design that utilizes the more familiar notion of a *uniform hazard*. Design for a targeted risk, *P[LS]*, can be achieved with the simple equation,

$$R_d = DF \cdot Q_d \tag{7}$$

[5] Earthquakes with a 2% probability of being exceeded in 50 years at a site (equivalently, a return period of 2,475 years) are referred to as the Maximum Considered Earthquake, or MCE, are mapped by the U.S. Geological Survey, and are the basis for earthquake-resistant structural design in ASCE Standard 7-05 (ASCE, 2005a).

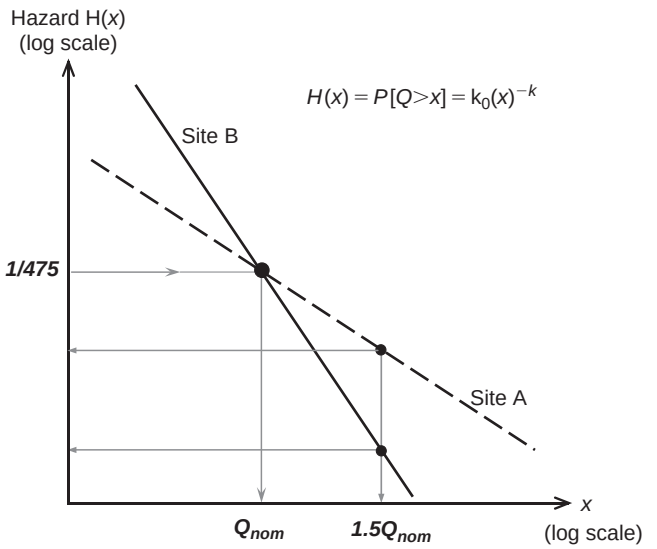


Figure 4 Design based on uniform hazard.

in which design factor, *DF*, is defined by:

$$DF = [\{H(R_d)/P[LS]\} \cdot \exp[(k\beta_R)^2/2 - 1.645\, k\beta_R]]^{1/k} \tag{8}$$

This approach has been adopted in the recently published *ASCE Standard 43-05* (ASCE, 2005b) on nuclear facilities.

To illustrate, suppose that one wishes to provide for "life safety under a rare earthquake event" (Figure 1) for a building. The U.S. Geological Survey furnishes uniform hazard maps for various probabilities of exceedence at its website (earthquake.usgs.gov/hazmaps/). The situation is depicted by Figure 4, which shows the hazard curves at two sites, A and B; at both sites, the seismic hazards are expressed by Eq (3) and the mapped value has a return period of 475 years. Assume that at site B, $k = 3.0$ (corresponding to a coefficient of variation of approximately 68% in the annual extreme demand). Such a value is typical in regions of high seismicity. The design goal is a structural system in which the probability of endangering life safety probability is less than 10% of the mapped uniform hazard exceedence probability. We adopt as our design criterion the requirement that the 5th percentile value of fragility determined from the system fragility assessment, R_d, must exceed the uniform hazard demand, *UHD*, multiplied by a design factor, *DF*. Assuming that the system fragility is characterized by a log-standard deviation, β_R, of approximately 0.50, which is a common value for steel and reinforced concrete frames (Ellingwood, Celik and Kinali, 2007) and using Eqs (7) and (8), the design requirement is that $R_d > 1.38\ UHD$. If site A is considered, $k = 1.5$, and the criterion would become $R_d > 2.46\ UHD$ Such a formulation provides a more systematic approach to ensuring performance than the

importance factor. In particular, if the (rare) hazard is mapped at 0.0021/yr (return period of 475 years), the limit state probability from this criterion would be 0.00021/yr at both sites.

As noted previously, the current performance objectives in PBE (summarized in Figure 1) are expressed in terms of uniform hazard (e.g., *life safety under a rare event*). In this case, the mapped value, Q_{nom}, is multiplied by a safety factor, say 1.5, against incipient collapse. It is clear from Figure 4 that the probability of exceeding $1.5Q_{nom}$ is vastly different for sites A and B. Using the same hazard and structural capacity parameters as in the previous paragraph, the incipient collapse probabilities at sites A and B are 0.0004/yr and 0.000016/yr, respectively, a difference of more than one order of magnitude. The uniform hazard approach fails to adequately reflect the effect of the slope of the hazard curve on risk. This problem is characteristic, in varying degrees, to all codes and standards with uniform hazard-based rather than uniform risk-based environmental load criteria.

4 Risk communication and competing risks

4.1 *Communicating risk*

Acceptable risks and target reliabilities are important ingredients of performance-based engineering. Selecting these benchmarks is difficult because of the small probabilities that often are involved, limitations in supporting data, and public attitudes toward risk. As decisions regarding performance and safety are increasingly made in the public arena, the issue of risk communication has become extremely important. Concepts of acceptable risk and what is acceptable building performance must be communicated in terms that are easily understood to the stakeholders of a building project.

The issue of what is acceptable risk in buildings was sidestepped in the development of the first-generation probability-based design standards, such as load and resistance factor design (LRFD), which are based on benchmark reliabilities of common member limit states (member yielding, buckling, plastic hinge formation) established by an assessment of traditional design practice (Ellingwood and Galambos 1982). No attempt was made to rationalize the calibrated reliabilities in explicit risk terms. The risks inherent to LRFD are related to social expectations of building system performance only to the extent that reliability benchmarks obtained from member calibration can be related to such expectations. Moreover, LRFD does not address system behavior explicitly (Ellingwood, 1994); rather, system effects are considered implicitly or are hidden in certain member limit states (e.g., through the use of K-factors in stability checks or R or C_d factors in earthquake-resistant design).

Human mortality from disease or accidents provides a natural starting point for benchmarking risks. Thresholds of risk acceptability are somewhat below the 10^{-3}/yr level associated with diseases (Reid, 2000) or roughly 10^{-4}/yr associated with motor vehicle accidents and home accidents. Comparisons of structural failure probabilities with such numbers are difficult because of differences in the population at risk, risk exposure, and the large epistemic uncertainties associated with structural reliabilities estimated from limited data. In the civil infrastructure area, the observation that, on average, 1% of dams fail (not necessarily catastrophically) within 20 years of construction implies a failure rate of about 5×10^{-4}/yr. Most estimates of building structural

collapse probability are on the order of 10^{-6}/yr to 10^{-5}/yr (Ellingwood, 2007). In first-generation probability-based codified design, the probability of yielding in a tension member or formation of a plastic hinge in a continuous ductile beam is on the order of 5×10^{-5}/yr (Ellingwood, 1994); the structural system collapse probability would be less, depending on the extent of load redistribution that is possible in the system subsequent to member failure. Thus, in a truss of 100 members, all of which were sized at the minimum design requirements, one might expect to see one or two members yield sometime during a service life of 50 years. These risks associated with current structural engineering, measured in terms of annual probability, are not inconsistent with acceptable risks for other involuntary social activities (on the order 10^{-7} to 10^{-5}/yr) (Ellingwood, 2007). On the other hand, such small probabilities are difficult for building stakeholders to interpret (Pate-Cornell 1994). Furthermore, any risk assessment that ignores the contribution of human error is likely to be skewed in the non-conservative direction.

It has become apparent that some decision-makers want a statement of confidence (as a measure of credibility) in the estimated risk on which the design decisions are based. This can take the form of a "confidence interval" [I am 95% confident that the losses will not exceed 5% of replacement cost] or a range [e.g., I expect the losses to be between 10% and 20% of replacement cost]. The first statement is understandable to a decision-maker trained in statistics. The second statement may be easier for one without technical training. Such measures can be developed by propagating the epistemic (knowledge-based) uncertainties through the risk analysis using simulation. An example of such an analysis in earthquake risk analysis can be found in Ellingwood, Celik and Kinali (2007). Most project stakeholders cannot distinguish between confidence levels of 90%, 95% or 99%; rather, they are "highly confident," "confident" or "indifferent," so a large number of confidence factors in such an analysis may not be necessary.

In the final analysis, determining acceptable risk hinges on socio-political as well as technical issues and concerns (Corotis, 2003). People are more willing to accept risks if they understand them. If quantitative risks or explicit numerical safety goals are to be used as a basis for structural codes, these concerns must be acknowledged and acceptable risks must be considered as outcomes of acceptable decision processes.

4.2 *Competing risks*

Performance-based engineering in the future will address newly identified hazards or hazards other than those for which the facility was designed. Design strategies for dealing with such issues can be addressed through the formalism of a probabilistic risk assessment (Faber and Stewart, 2003), a technology which has become feasible for certain civil infrastructure projects. A basic mathematical framework for building risk assessment that can be used for this purpose is provided by (Ellingwood, 2007):

$$P[Loss > \vartheta] = \sum_{H} \sum_{LS} \sum_{DS} P[Loss > \vartheta | DS] \cdot P[DS|LS] \cdot P[LS|H] \cdot P[H] \qquad (9)$$

in which the term $P[H]$ is annual probability of occurrence of hazard H (for rare events, $P[H]$ is numerically equivalent to the annual mean rate of occurrence, λ_H,

which is more easily estimated from available data); $P[LS|H]$ = conditional probability of a structural limit state (yielding, instability), given the occurrence of H; $P[DS|LS]$ = conditional probability of damage state DS (e.g., negligible, minor, moderate, major, severe) arising from structural damage, and $P[Loss > \vartheta | \ DS]$ = annual probability (mean frequency) of loss exceeding ϑ, given a particular damage state. The loss metric, $P[Loss > \vartheta]$, involves number of injuries or deaths, damage costs exceeding a fraction of overall replacement costs, loss of opportunity costs, etc, depending on the objectives of the assessment. Although the event $\{Loss > \vartheta\}$ usually has been replaced with {Life-threatening damage} or {Collapse} for buildings, bridges and other civil infrastructure because of concerns about life safety, the need to minimize or control direct and indirect economic losses might cause other metrics to be added to this traditional metric in the future.

Eq (6) deconstructs the risk analysis into its major constituents and show that appropriate strategies for risk mitigation should be a multidisciplinary effort aimed at three basic levels: (i) to prevent the occurrence of hazardous events through social or political means; (ii) to prevent the occurrence of local significant structural damage that is likely to initiate a collapse; and (iii) to prevent structural system collapse and loss of life through innovative structural design. All sources of uncertainty, from the hazard occurrence to the response of the structural system, must be considered and propagated through the risk analysis framework defined by Eq (9). The structural reliability methods described in detail elsewhere in this paper are essential to achieving the goal of risk mitigation but provide a more limited perspective, addressing only the final two terms in Eq (9). Nonetheless, structural reliability provides the starting point for PBE and for risk-informed decision-making.

5 Conclusions

Performance-based engineering has the potential to clarify the objectives of design, to facilitate understanding of the code and the intent of its provisions, and to enables structural engineers to devise alternative solutions that meet performance expectations equally well. This encourages the development of innovative design solutions and enhances beneficial competition in the building community. Moreover, it provides a rational framework for evaluating existing structures, where the characteristics of available information and data have a different character than in new construction, and thus is an important tool for civil infrastructure renewal. PBE enables the building owner or occupant (consumer) to manage risk by increasing his investment in additional building strength and stiffness or in uncertainty reduction. With performance goals expressed for different categories of construction in probabilistic terms, modern decision analysis tools that are based on risk can better support post-disaster response, assessment and rehabilitation of facilities, insurance underwriting, resource allocation, and public policy. Better understanding of the relation between building design, behavior and performance, coupled with uncertainty/sensitivity analysis, will lead to additional research that is focused on topics that will enhance future performance. Perhaps most important, PBE provides a communications interface between building owners and occupants, engineers and architects, thereby facilitating the professional practice of structural engineering.

Acknowledgement

The views expressed in this paper are based on research supported for many years by several organizations, including the US Nuclear Regulatory Commission, Oak Ridge National Laboratory, the US Army Corps of Engineers, the National Science Foundation, and the National Institute of Standards and Technology. This support is gratefully acknowledged. However the opinions expressed herein are solely those of the writer, and do not necessarily represent the views of the sponsoring organizations.

References

ASCE/ANSI Standard 7-05. *Minimum design loads for buildings and other structures*. American Society of Civil Engineers, Reston, VA, 2005a.

ASCE Standard 43-05, *Seismic design criteria for structures, systems and components in nuclear facilities*. American Society of Civil Engineers, Reston, VA, 2005b.

ASCE/ANSI Standard 41-06. *Guidelines for the Seismic Rehabilitation of Buildings* (previously published as FEMA 356), American Society of Civil Engineers, Reston, VA, 2006.

Cornell, C.A., Jalayer, F., Hamburger, R.O. and Foutch, D.A. Probabilistic basis for 2000 SAC FEMA steel moment frame guidelines. *Journal of Structural Engineering ASCE*, Vol. 128, No. 4, pp. 526–533, 2002.

Corotis, R.B. Socially relevant structural safety. *Applications of Statistics and Probability in Civil Engineering, Proc. ICASP 9* (Der Kiureghian, Madanat and Pestana, eds.), Millpress, Rotterdam, pp. 15–26, 2003.

Ellingwood, B.R. and Galambos, T.V. Probability-Based Criteria for Structural Design. *Structural Safety*, Vol. 1, No. 1, pp. 15–26, 1982.

Ellingwood, B.R. Probability-based codified design: past accomplishments and future challenges. *Structural Safety*, Vol. 13, No. 3, pp. 159–176, 1994.

Ellingwood, B.R. Reliability-based performance concept for building construction in *Structural Engineering Worldwide 1998*, Paper T161-4, Elsevier Science Ltd, 1998.

Ellingwood, B.R. Strategies for mitigating risk to buildings from abnormal load events. *International Journal of Risk Assessment and Management,* Vol. 7, Nos. 6/7, pp. 828–845, 2007.

Ellingwood, B.R., Celik, O.C. and Kinali, K. Fragility assessment of building structural systems in mid-America. *Earthquake Engineering & Structural Dynamics* Vol. 36, pp. 1935–1952, 2007.

Faber, M. and Stewart, M.G. Risk assessment for civil engineering facilities: a critical appraisal. *Reliability Engrg. & System Safety*, Vol. 80, pp. 173–184, 2003.

Hamburger, R.O., Foutch, D.A. and Cornell, C.A. Translating research to practice: FEMA/SAC performance-based design procedures. *Earthquake Spectra*, Vol. 19, No. 2, pp. 255–267, 2003.

Melchers, R.E. *Structural reliability: analysis and prediction*. John Wiley, Chichester, UK, 1999.

Pate-Cornell, E. Quantitative safety goals for risk management of industrial facilities. *Structural Safety*, Vol. 13, No. 3, pp. 145–157, 1994.

Reid, S.G. Acceptable risk criteria. *Progress in Structural Engineering and Materials* Vol. 2, No. 4, pp. 254–262, 2007.

Singhal, A. and Kiremidjian, A.S. Method for probabilistic evaluation of seismic structural damage. *Journal of Structural Engineering, ASCE*, Vol. 122, No. 12, pp. 1459–1467, 1996.

Shinozuka, M., Feng, M.Q., Lee, J. and Naganuma, T. Statistical analysis of fragility curves. *Journal of Engineering Mechanics, ASCE*, Vol.126, No. 12, pp. 1224–1231, 2000.

Wen, Y. K. and Ellingwood, B. R. The role of fragility assessment in consequence-based engineering. *Earthquake Spectra* EERI, Vol. 21, No. 3, pp. 861–877, 2005.

Chapter 9

Bayesian network methodology for post-earthquake infrastructure risk management

Armen Der Kiureghian, Michelle Bensi & Daniel Straub

University of California, Berkeley, CA, USA

ABSTRACT: In the immediate aftermath of an earthquake affecting infrastructure systems, decisions must be made regarding the deployment of emergency personnel and equipment, evacuation of people, inspection, closure or opening of facilities, and other actions to assure safety of people. Furthermore, soon after the earthquake event, selections must be made among alternative actions to restore functionality to vital infrastructure services. The key ingredient for such decision-making is information: information about the nature and characteristics of the hazard, about the states of the system and its components, and about the consequences of various decision alternatives. In the aftermath of an earthquake, the available information is usually incomplete, highly uncertain, and rapidly evolving in time. We present a Bayesian network methodology for information updating regarding an infrastructure subject to earthquake hazard. Given observed information about the earthquake and the infrastructure components, such as sensor measurements of the ground motion at selected points or observations of damage/no damage of the infrastructure components, we use Bayesian updating to assess the probabilistic state of the infrastructure both at the local (component) and global (system) levels. This analysis properly accounts for the spatial correlation structure of earthquake ground motions, which has been neglected in previous studies.

1 Introduction

Infrastructures, such as transportation, water and power networks, are backbones of modern societies. Their resilience in face of natural and man-made hazards is vital for the well-being of communities. Earthquakes are a dominant hazard to these systems in many geographic regions of the world. In the immediate aftermath of an earthquake affecting an infrastructure, decisions must be made regarding the deployment of emergency personnel and equipment, evacuation of people, inspection, closure or opening of facilities, and other actions to assure safety of people. Furthermore, soon after the earthquake, selections must be made among alternative actions to restore functionality to vital infrastructure services. The key ingredient for such decision-making in face of a hazard is information: information about the nature and characteristics of the earthquake, about the states of the system and its components, and about the consequences of various decision alternatives. However, in the chaotic aftermath of an earthquake, the available information is usually incomplete, highly uncertain, and rapidly evolving in time.

In this paper we present a Bayesian network methodology for information updating regarding an infrastructure subject to earthquake hazard. Given observed information about the earthquake and the infrastructure components, such as sensor measurements of the ground motion at selected points or observations of damage/no damage of the infrastructure components, we use Bayesian updating to assess the probabilistic state of the infrastructure both at the local (component) and global (system) levels. This analysis properly accounts for the spatial correlation structure of earthquake ground motions, which has been neglected in many previous studies. By extending the Bayesian network with utility and decision nodes, it is possible to provide a preference ordering of the action alternatives that may be available to individuals responsible for decision-making. The eventual goal of this research is to develop a near-real time decision-support system for emergency response, recovery and risk management of infrastructures subject to seismic events.

A near real-time decision-support system for an infrastructure subject to a hazard must address a number of key issues. These include: (a) proper characterization of the infrastructure as a system, including interactions and dependencies among its components as well as dependence with other infrastructures (e.g., the dependence of water and power systems); (b) proper account of the uncertainties that are present in the characterization of the hazard and the states of the system components; (c) ability to update the characterization of the hazard and the probabilistically assessed states of the infrastructure components as information becomes available (i.e., reduce uncertainties); (d) computational efficiency and scalability in order to allow near real-time application to large and complex systems; and (e) the provision of a user-friendly interface not only for assistance in decision-making, but also in developing the infrastructure model.

One objective of this paper is to introduce and promote the Bayesian network (BN) methodology as one that possesses all the key requirements mentioned in the preceding paragraph. Specifically, the BN facilitates probabilistic modeling of systems involving a large number of statistically dependent random variables and it enables efficient updating of the random variables and the system state given a set of observations. Furthermore, the BN has a graphical nature, which facilitates its use by modelers and decision-makers. On the other hand, the BN has certain limitations in terms of the type of statistical dependence that it can handle. For example, analysis with a large number of random variables drawn from a random field, e.g., ground motion intensities at critical locations of a spatially distributed infrastructure, may require enormous computational effort, making near real-time applications impractical. This limitation and a corresponding approximate solution approach are described in this paper.

The paper begins with a brief introduction to the BN methodology. This is followed by the development of a BN model for Bayesian seismic hazard analysis of a general spatially distributed system. The model highlights the need to address random fields of ground motions and the difficulty that this causes in the computations with the BN. An approximation method based on principal component analysis is then introduced to simplify the BN model and reduce its computations. This methodology is then applied to an example transportation system with the aim of demonstrating the BN methodology as well as the accuracy of the truncation method based on the principal component analysis. The states of selected components as well as the highway system are updated for various types of observations related to the ground motion or the components of the system.

The research described in this paper is in its early stage. Therefore, the reported results demonstrate only a fraction of the potential that the BN methodology offers for application to infrastructure risk analysis and management. We hope our future publications will provide broader and more in-depth developments and results.

2 Brief on Bayesian network

A BN is a graphical representation of a set of dependent random variables. The variables may describe quantities affecting a system, or they may describe the states of the system components or the system itself. The BN is a directed acyclic graph (DAG), in which nodes represent random variables and directed links represent statistical dependence. To satisfy rules of probability, the graph cannot contain closed loops – hence the requirement that the graph be acyclic.

Figure 1 shows a simple BN with five random variables, $X_1, \ldots, X_5$. As indicated by the directed links, X_3 is defined conditional on X_1 and X_2, X_4 is defined conditional on X_1, and X_5 is defined conditional on X_4. We say X_3 is a child of X_1 and X_2, while the latter are the parents of X_3, etc. Considering the dependence structure encoded in the BN graph, the joint probability mass function of the five variables can be written as

$$p(x_1, x_2, x_3, x_4, x_5) = p(x_5|x_4)p(x_4|x_1)p(x_3|x_1, x_2)p(x_1)p(x_2) \tag{1}$$

The BN facilitates probabilistic analysis by recognizing the dependence structure among the random variables. To appreciate this fact, assume each of the five random variables in Figure 1 has m possible states (possible outcomes). To directly specify the joint distribution in the left-hand side of (1), we would need to specify $m^5 - 1$ values (minus one because the sum of all probabilities should add up to 1). On the other hand, the number of probability terms required to specify the right-hand side is $2m^2 + m^3 + 2m - 5$. For large m, the specification in terms of conditional probabilities is far more efficient (by two orders of magnitude in the present case). Therein lies the efficiency of the BN.

Generalizing (1), the joint distribution of random variables in a BN can be written as

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i|pa(X_i)) \tag{2}$$

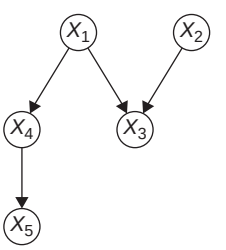


Figure 1 A simple BN.

where $pa(X_i)$ denotes the parents of random variable X_i. To complete the BN, to each node X_i is attached a conditional probability table (called potential) listing the probability values $p(x_i|pa(X_i))$. The size of this table clearly depends on the number of parents of X_i and the number of states of X_i and each of its parents. It should be clear, therefore, that a node with many links converging into it will require a high-dimensional probability table. Computations involved in assessing the marginal distributions of the nodes as well as updated distributions are also strongly affected by the number of links converging into a node. For efficient computation, it is desirable that each BN node have no more than 3 or 4 links directed towards it. This number could be higher if the random variables have small numbers of states.

The most important advantage of the BN is the ability to update probability distributions of the nodes when "evidence" is entered at one or more nodes. By evidence, we mean information regarding the outcomes of the random variables. This is shown in Figure 2, where evidences are entered at nodes X_3 and X_5. These could be observed outcomes of these random variables. Given such evidence, one might be interested in updated probability distributions, such as $p(x_1|e_3, e_5)$ or $p(x_1, x_2|e_3, e_5)$. This updating takes advantage of the so-called d-separation rules (Jensen 2001). Observe in Figure 1 that connections in BN can be of three types: serial connection, as that of X_1, X_4 and X_5; converging connection, as that of X_1, X_2 and X_3; and diverging connection, as that of X_1, X_3 and X_4. One can show that evidence transmits through a serial connection, unless the middle node is known with certainty; evidence transmits through a diverging connection, unless the parent node is known with certainty; and evidence transmits through a converging connection if there is some evidence for the child node or one of its descendants. These properties are used to develop efficient algorithms for transmitting information through a BN to obtain updated probability distributions of the nodes. More details on BN and its various computational algorithms can be found in Pearl (1988), Jensen (2001) and Langseth and Portinale (2007). Applications of BN to systems reliability problems can be found in, e.g., Weber and Jouffe (2006) and Wilson et al. (2007).

In the foregoing, the random variables X_i have been assumed to be discrete with a finite number of possible outcomes. Although certain continuous random variables can be handled by the BN, in practice most such variables will need to be discretized, especially if exact inference algorithms are to be used. This constitutes one shortcoming of the BN. Yet, experience shows that for most risk analysis problems effective discretizations can be used with little loss of accuracy (see, e.g., Friis-Hansen 2005).

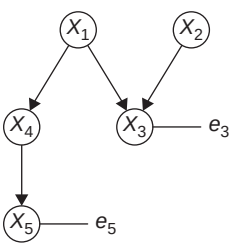


Figure 2 BN with evidence.

3 Bayesian probabilistic seismic hazard analysis

Probabilistic seismic hazard analysis (PSHA) involves assessment of the probability that a ground motion intensity measure will exceed a specified threshold at one or more locations (Cornell 1968). For a spatially distributed system, the joint distribution of intensity measures at several locations is of interest. More specifically, let S_i denote the intensity measure, e.g., the spectral acceleration at the natural period of a critical infrastructure component, at location i. Marginal seismic hazard analysis involves determination of $\Pr(s_i < S_i)$. If the component has capacity s_i, then the above describes the probability of failure of the component in case of the earthquake. For a system, let C_k denote the k-th minimal cut set, i.e., the k-th minimal set of components whose joint failure constitutes failure of the system. Then, the probability of failure of the system is described by $\Pr(\bigcup_k \bigcap_{i \in C_k} s_i < S_i)$, where the intersection is over all the elements within the k-th minimal cut set and the union is over all the minimal cut sets. It should be clear that the joint distribution of the S_i's is necessary in order to compute this probability.

In this paper, following Straub et al. (2008), we present a computational framework that enables assessment of the conditional probability $\Pr(\bigcup_k \bigcap_{i \in C_k} s_i < S_i | E)$, where E is any available evidence (observations) related to the seismic hazard and/or the infrastructure performance. Such observations may include measurement of earthquake magnitude and location, measurement of ground motion intensities at selected sites, and observations of performances of various components of the infrastructure system. In deference to the Bayes' rule for probability updating, we designate this approach as Bayesian PSHA.

The BN modeling technique is particularly well suited for Bayesian PSHA due to its computational efficiency in performing probability updating in light of new evidence. Figure 3 sketches the proposed Bayesian network for an infrastructure system. For simplicity, the BN in Figure 3 is developed to represent only a single earthquake event. In this paper, we limit ourselves to this case, but the framework can be extended to the case of multiple earthquakes randomly occurring in time.

As can be seen in Figure 3, the top node of the BN defines the occurrence or non-occurrence of the earthquake event. The earthquake magnitude and location are directly dependent on the event of occurrence. The intensities S_i at n locations of the infrastructure system are dependent on the earthquake magnitude and distance. In addition, these variables are correlated with each other, as the directed links connecting them indicate. The source of this correlation is described below and in the next section. The nodes D_1 to D_n describe the states of the components of the infrastructure system at the n locations. These directly depend on the corresponding ground motion intensities, but also on the capacity of each component, which can be random. The probability distribution of each node D_i conditioned on the corresponding S_i is defined by the fragility function of the component. Finally, the system performance node describes the state of the system in terms of the states of the components. This relationship usually is deterministic. That is, given the states of the components, one determines if any of the minimal cut sets of the system has been realized, in which case the system is in the fail state; otherwise, the system is in the survival state.

Seismic intensities S_i at different sites i are computed using a ground motion prediction model in terms of the earthquake magnitude, m, and the distance of the site from the earthquake source, r_i. This relationship is typically developed by regression

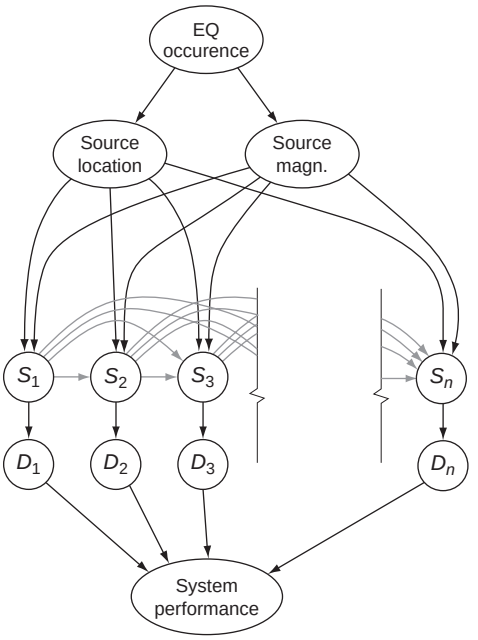


Figure 3 BN for an infrastructure system subject to earthquake hazard.

analysis of recorded data and contains a random error term, so that S_i is a random variable even for given m and r_i. Due to common influencing factors, the error terms for different sites, and consequently the S_i, are correlated (Min and Takada 2005). In Figure 3, this dependence is represented by the links among the S_i variables. This model leads to $n+1$ links pointing towards the last node, S_n. (In the BN terminology, variable S_n has $n+1$ parents.) As discussed earlier, when applying exact inference algorithms in the BN, the number of parents to a single variable is critical, since it directly influences the computation effort required for inference. The computation time and required computer memory grows approximately exponentially with the number of parents, thus making the BN formulation infeasible for modeling systems with more than a few components. This limitation motivates the introduction of an approximate model of statistical dependence in space, which is introduced in the following section.

4 Approximate model of spatially correlated ground motion intensities

For an earthquake of given magnitude and location, let the intensity S_i at a site i be described by a ground motion prediction model of the form

$$\ln(S_i) = g(m, r_i, \boldsymbol{\theta}_i) + \varepsilon_m + \varepsilon_{r,i} \tag{3}$$

where $g(m, r_i, \boldsymbol{\theta}_i)$ is the deterministic prediction relation as a function of the magnitude m, the distance of site i from the earthquake source, r_i, and parameters $\boldsymbol{\theta}_i$ specifying

the local site conditions and structural properties, e.g., the fundamental period for determining spectral acceleration values. ε_m and $\varepsilon_{r,i}$ in (3) are zero-mean normal random variables representing the inter-event and intra-event error terms, respectively, as obtained from the regression of recorded data (Park et al. 2007). Let σ_m and σ_r denote the standard deviations of ε_m and $\varepsilon_{r,i}$, respectively, and $\rho_{r,ij}$ denote the correlation coefficient between pairs of $\varepsilon_{r,i}$ and $\varepsilon_{r,j}$. The latter is available from the specification of an autocorrelation function for the random field of ground motion intensities, see (Min and Takada 2005, Park et al. 2007). Thus, for given m and r_i, the set of log-intensities $\ln(S_i)$, $i = 1, \ldots, n$, are jointly Normal with means $\lambda_i = g(m, r_i, \boldsymbol{\theta}_i)$, common variance $\zeta^2 = \sigma_m^2 + \sigma_r^2$, and correlation coefficients

$$\rho_{ij} = \rho_{\ln S_i, \ln S_j} = \frac{\sigma_m^2 + \sigma_r^2 \rho_{r,ij}}{\sigma_m^2 + \sigma_r^2} \tag{4}$$

For ease of notation, we define $\varepsilon_i = \ln(S_i) - \lambda_i = \varepsilon_m + \varepsilon_{r,i}$ as the total error. Thus, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots \varepsilon_n)$ has the joint normal distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}$, whose elements are $\zeta^2 \rho_{ij}$.

To facilitate a computationally efficient representation of $\boldsymbol{\varepsilon}$ in a BN, we introduce the transformation

$$\boldsymbol{\varepsilon} = \mathbf{TU} \tag{5}$$

where $\mathbf{U} = [U_1, U_2, \ldots, U_n]^{\mathrm{T}}$ is a vector of statistically independent standard normal random variables and $\mathbf{T}$ is an $n \times n$ transformation matrix. We require that $\mathbf{T}$ be the Karhunen-Loève transformation of $\boldsymbol{\varepsilon}$, as obtained from the principal component analysis (PCA) (Jolliffe 2002). Therefore, $\mathbf{T} = \boldsymbol{\Phi}\mathbf{V}^{1/2}$, where $\boldsymbol{\Phi} = [\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_n]$ is obtained by solving the eigenvalue problem

$$\boldsymbol{\Sigma}\boldsymbol{\Phi}_i = \boldsymbol{\Phi}_i \upsilon_i, \quad i = 1, \ldots, n \tag{6}$$

with $\mathbf{V} = \mathrm{diag}[\upsilon_i]$ being the diagonal matrix of eigenvalues. We order the eigenvalue/eigenvector pairs $(\upsilon_i, \boldsymbol{\Phi}_i)$ such that $\upsilon_1 > \upsilon_2 > \cdots > \upsilon_n$. The vector $\boldsymbol{\varepsilon}$ is now approximated by the vector $\boldsymbol{\varepsilon}'$ defined as

$$\boldsymbol{\varepsilon}' = \boldsymbol{\Phi}'\mathbf{V}^{1/2}\mathbf{U} + \mathbf{K} \tag{7}$$

where $\boldsymbol{\Phi}'$ is equal to $\boldsymbol{\Phi}$ with a number of elements replaced by the value zero, and $\mathbf{K}$ is a vector of zero mean, statistically independent normal random variables that ensures that the variances of $\boldsymbol{\varepsilon}'$ are equal to the variances of $\boldsymbol{\varepsilon}$. One can easily show that the correction terms have the variances

$$\sigma_{K_i}^2 = \zeta^2 - \left[\boldsymbol{\Phi}'\mathbf{V}\boldsymbol{\Phi}'^{\mathrm{T}}\right]_{i,i} \tag{8}$$

The approximation of $\boldsymbol{\varepsilon}$ by $\boldsymbol{\varepsilon}'$ in the BN can be interpreted as follows: First, consider the transformation in (4), which is included in the BN by introducing the variables $\mathbf{U}$ with links from all $\mathbf{U}$ to all $\boldsymbol{\varepsilon}$ (Figure 4a). The dependence among the $\boldsymbol{\varepsilon}$ is then represented by the common parents $\mathbf{U}$. Setting an element $[\boldsymbol{\Phi}]_{ij} = 0$ corresponds to eliminating the link from U_j to ε_i in the BN. To enable efficient computation of the BN, sufficient

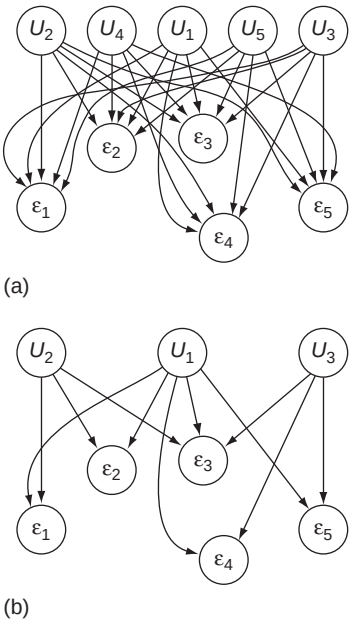


Figure 4 Elimination of links.

links must be eliminated. This approximate BN model is illustrated in Figure 4b. To determine which links to eliminate, it is proposed to establish a marginal importance measure of the links as

$$M_{ij} = \sum_{k,l} \Delta_{kl} \tag{9a}$$

where Δ_{kl} are the elements of the matrix

$$\mathbf{\Delta} = \left| \mathbf{\Sigma} - \mathbf{\Phi}'_{ij} \mathbf{V} \mathbf{\Phi}'^{\mathrm{T}}_{ij} \right| \tag{9b}$$

wherein $\mathbf{\Phi}'_{ij}$ is the matrix $\mathbf{\Phi}$ with element $[\mathbf{\Phi}]_{ij}$ set to zero. The links are now ordered according to M_{ij} and links with lower marginal importance measures are eliminated first. After elimination of links, if a variable U_i has zero remaining links, the corresponding node can be eliminated. Furthermore, if U_i has a single remaining link, it can also be eliminated without further loss of accuracy. Finally, the conditional distribution of $\boldsymbol{\varepsilon}'$ given $\mathbf{U} = \mathbf{u}$ is Gaussian with mean vector $\mathbf{M}_{\boldsymbol{\varepsilon}'|\mathbf{u}} = \mathbf{\Phi}' \mathbf{V}^{1/2} \mathbf{u}$ and covariance matrix $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}'|\mathbf{u}} = \mathrm{diag}[\sigma^2_{K_i}]$. These values are used to construct conditional probability tables for the nodes ε_i in the reduced BN in Figure 4b. Implementing this approximation in the original BN model of Figure 3, the final BN representation with reduced links is shown in Figure 5.

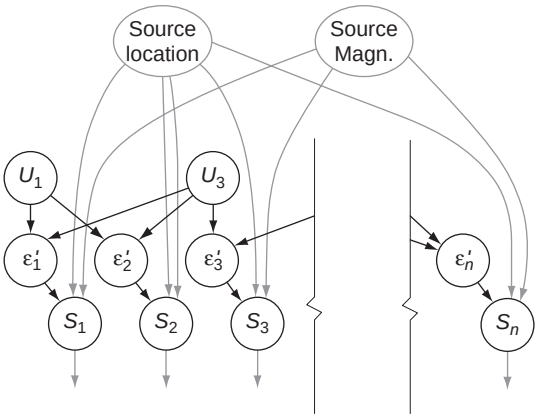


Figure 5 Representation of ground motion intensities in the reduced BN.

5 BN computations

In the numerical example presented in the remainder of this paper, all continuous random variables are discretized to enable exact inference in the BN, which is desirable because of computational efficiency. Most variables are discretized at 10 intervals. The discretization scheme has been successfully tested against simulation results, which are readily obtained for the unconditional (no-evidence) case. It is noted that approximate inference algorithms that can handle continuous random variables exist, such as the Gibbs sampler (Gilks 1996). The disadvantage of these algorithms is their unknown rate of convergence. Exact inference algorithms are therefore deemed preferable for applications of the model in near-real time decision applications. However, the application of approximate algorithms should be studied further in the future.

6 Numerical investigation

In the remainder of the paper, the proposed model is investigated by application to a transportation system, which is based on an example from Kang et al. (2007). The system consists of 12 bridges, which are part of a network connecting a number of cities to a hospital. The geographical location of the bridge sites is shown in Figure 6. The fragility model of the 12 bridges is a modified version of that used by Kang et al. (2007). The earthquake hazard is a point source along a fault line, which starts at ($x = 0$, $y = 0$) and extends for 50 km in x-direction. The source is assumed to occur with equal likelihood anywhere along the fault line. The PDF of the magnitude and the ground motion prediction model are as in Kang et al. (2007). The standard deviations of the error terms ε_m and $\varepsilon_{r,i}$ are $\sigma_m = 0.1$ and $\sigma_r = 0.5$. Spatial correlation among the $\varepsilon_{r,i}$ is described by the isotropic autocorrelation function proposed by Boore (2003) as a function of the distance r_{ij} between two sites,

$$\rho_{r,ij} = 1 - \left[1 - \exp\left(-\sqrt{r_{ij} 0.6\,\mathrm{km}^{-1}}\right)\right]^2 \tag{10}$$

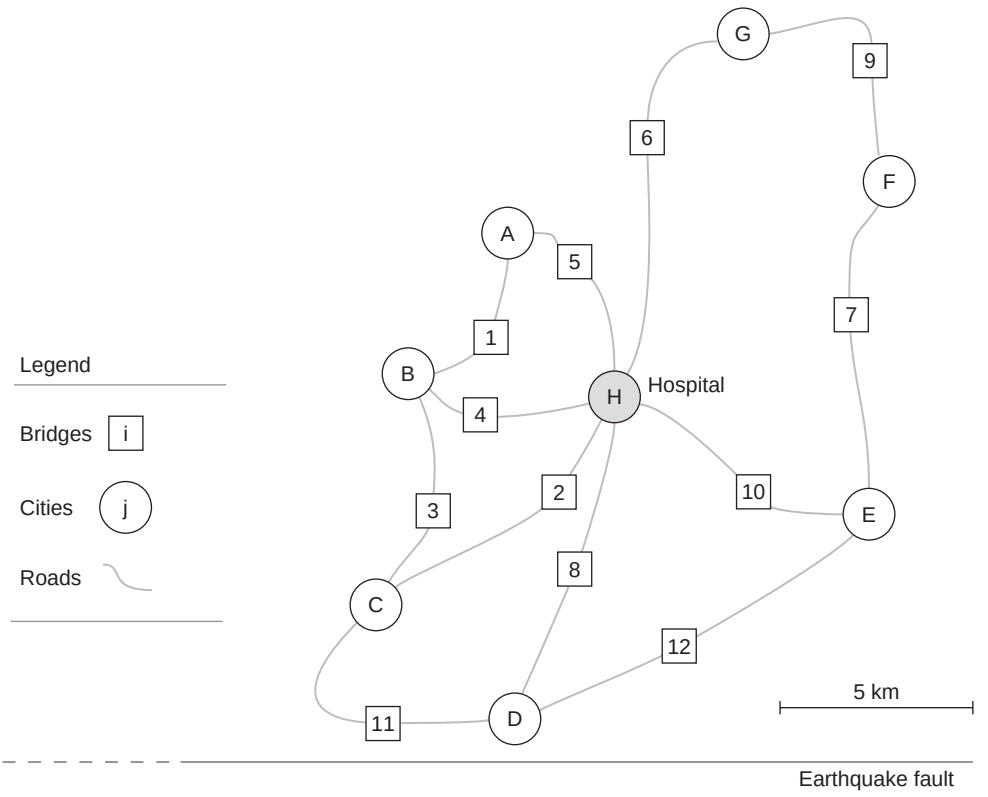


Figure 6 The geographic layout of the example infrastructure system.

6.1 *Information updating*

We first investigate the influence of information on the estimation of the seismic intensity at the site of bridge 1 (Site 1) during an earthquake event.

Shown in Figure 7 are three probability distributions of the ground motion intensity at Site 1, as measured in terms of the peak ground acceleration, PGA. The dotted curve represents the distribution of the intensity when no information other than the occurrence of the earthquake is available. Observe that the likelihood is small that the intensity will be large. This is due to the small likelihood of a large magnitude earthquake (relative to small magnitude events). The other two lines show the distribution of the intensity at Site 1 when a PGA of 0.75 g at Site 4 has been measured. The dashed line is for the case where the spatial correlation in the error terms is neglected, while the solid line is for the case where the spatial correlation is accounted for. In the former case, the effect of the evidence comes through the information that the measured PGA at Site 4 provides about the magnitude and location of the earthquake. In the latter case, where the effect of the positive spatial correlation between the errors is included, the observation provides additional information about the PDF of the PGA at site 1, resulting in a further shift of the PDF towards higher values. Observe

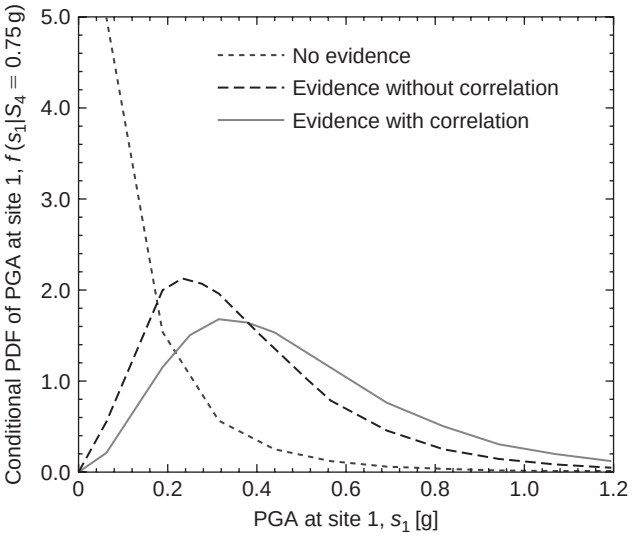


Figure 7 Updated distribution of ground motion intensity at Site 1.

that the information provided by the measurement at Site 4 profoundly influences the distribution of the ground motion intensity at Site 1.

The computation time required to perform inference in the BN for the presented example is in the order of 60 CPU seconds on a standard Pentium II computer. These computation times were achieved with a generic BN inference algorithm (modified Pearl algorithm). We have observed that this algorithm is not optimal for the proposed BN model. To significantly improve the computational performance (by orders of magnitude), a custom-made inference algorithm is being developed, which can make use of the problem-specific dependence structure.

6.2 *Investigation of the approximate model*

Figure 8 demonstrates the effect of the approximate model proposed in this paper for the instantiation of one set of evidence, namely $m = 6$, source at 25 km from the left end of the fault, and a measured PGA of 0.75 g at Site 4. Different degrees of approximation are characterized in terms of the number of links from nodes U_j to nodes ε_i that are retained in the model. The links are selected according to their marginal importance measures as in (9). The computations of the updated complementary CDF in Figure 8 are performed using stochastic simulation.

The results in Figure 8 show that the approximate model represents an improvement over neglecting the correlation in the error terms. However, a disadvantage of the approach is that it is not possible, in general, to predict if the approach under- or overestimates the quantity of interest. For the example presented in Figure 8, the model with 20 links leads to an unconservative estimate, whereas the model with 40 links is conservative; however this trend was not consistent for the wide range of

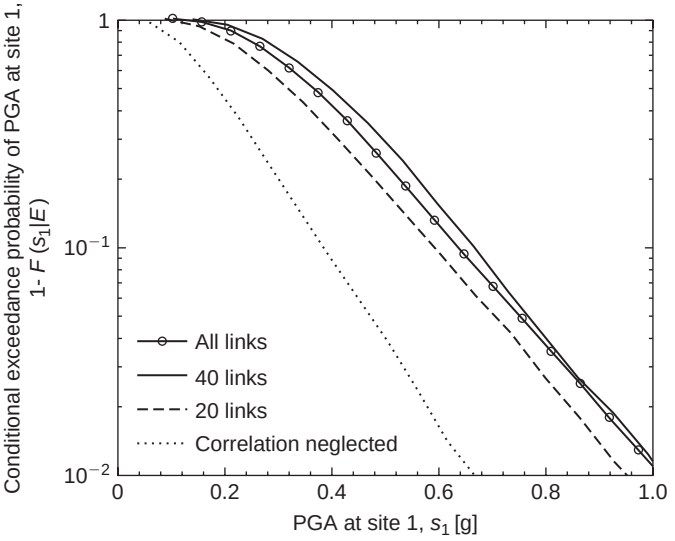


Figure 8 Influence of link elimination.

evidence cases examined but not presented in this paper. This remains a topic for future investigation.

6.3 *System reliability*

Inclusion of the error terms and account of the spatial correlation among the error terms have significant influences on the system reliability. Here, we are interested in the connectivity between city A and the hospital H, as shown in Figure 6. The corresponding system failure probability Pr(F) for different evidence instantiation cases are listed in Table 1. Included are system failure probability estimates without any evidence (first three rows), but with different treatment of the error terms. It can be seen that neglecting the error terms underestimates the failure probability by more than one order of magnitude; furthermore, neglecting the spatial correlation of the error terms results in an underestimation of the risk by almost a factor of 3. The failure probability estimate drastically increases when a PGA of 0.75 g is measured at Site 4 (rows 4 and 5), or when the bridge at Site 9 is observed to have failed (rows 6 and 7). In these cases, the reduced model with only 40 links provides fairly good approximation when compared with the exact model with all 140 links.

The results in Table 1 were obtained using a simulation-based (approximate) BN inference algorithm. These computations can become increasingly demanding, in particular for small probabilities of failure. Here probabilities are computed with 10^5 samples. The CPU time is in the order of 10 seconds, but the computational effort can increase significantly depending on the evidence and the applied algorithm.

The results in Table 1 demonstrate the importance of including the full error model in the analysis. Because of the redundant nature of the network system, neglecting the correlation leads to an underestimation of the system failure probability. Table 1

Table 1 System failure probability Pr(*F*) for different cases.

Case	*Pr(F)*
Unconditional (neglecting error terms and correlation)	0.00009
Unconditional (neglecting error correlation)	0.0012
Unconditional (full model: including error terms and correlation)	0.0028
Conditional on observed PGA at site 4, PGA(4) = 0.75 g (full model, all links)	0.046
Conditional on observed PGA at site 4, PGA(4) = 0.75 g (full model, 40 links)	0.050
Conditional on failure of bridge 9 (full model, all links)	0.054
Conditional on failure of bridge 9 (full model, 40 links)	0.060

also illustrates the effect of including additional evidence regarding seismic intensities and site failure/survival information on the calculation of system failure probabilities. Specifically, Table 1 shows pronounced effect of observing a high PGA value at a site or the effect of a bridge failure, even when the bridge is not located on a direct path from city A to the hospital.

7 Conclusion

The relevance and usefulness of the Bayesian network methodology for infrastructure risk management in face of seismic hazard is demonstrated. The methodology can be used to update probabilistic information about the state of an infrastructure system and its components as information from measurements or other observed states of system components are received. Such updated information is valuable in supporting critical decisions that need to be made in the aftermath of a damaging earthquake. The specific application considered demonstrates the need for a model that includes the spatial correlation among the seismic intensities at different locations of a spatially distributed system. The proposed reduction method allows accounting of the spatial correlation within the BN framework without unduly increasing the computational effort. The model is computationally efficient, and thus suitable for application in near real-time decision-making. As an example, the model can be used to provide estimates of seismic intensities following an earthquake, which can help in prioritizing emergency response and recovery actions.

Acknowledgment

The second author is supported by a National Science Foundation Graduate Research Fellowship. The third author acknowledges support by the Swiss National Science Foundation through grant PA002-111428. Additional support from the Taisei Chair Fund at the University of California, Berkeley is acknowledged.

References

Boore, D.M., et al. Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard Bridge collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America*, **93**:2737–2751, 2003.

Cornell, C.A. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, **58**:1583–1606, 1968.

Friis-Hansen, P. Structuring of complex systems using Bayesian network. *Proceedings Workshop on Reliability Analysis of Complex Systems*, Technical University of Denmark, Lyngby, August 2004, pp. 111–133, 2005.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter. *Markov chain Monte Carlo in practice*. Chapman & Hall, New York, NY, 1996.

Jensen, F.V. *Bayesian networks and decision graphs*. Springer, New York, 2001.

Jolliffe, I.T. *Principal Component Analysis*. 2nd ed. Springer Series in Statistics. Springer, New York, NY, 2002.

Kang, W.-H., J. Song, and P. Gardoni. Matrix-based system reliability method and applications to bridge networks. *Reliability Engineering & System Safety*, (accepted for publication) 2007.

Langseth, H. and L. Portinale. Bayesian networks in reliability. *Reliability Engineering & System Safety*, **92**:92–108, 2007.

Min, W. and T. Takada. Macrospatial correlation model of seismic ground motions. *Earthquake Spectra*, **21**:1137–1156, 2005.

Park, J., P. Bazzurro, and J.W. Baker. Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. in *Proceedings, 10th International Conference on Applications of Statistics and Probability in Civil Engineering*. Tokyo, Japan, Taylor & Francis Group, London, UK, 2007.

Pearl, J., *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers, San Francisco, CA, 1988.

Straub, D., M. Bensi and A. Der Kiureghian (2008). Spatial modeling of earthquake hazard and infrastructure performance through Bayesian Network. *Proceedings, Inaugural International Conference of the Engineering Mechanics Institute, ASCE*, Minneapolis, MN, May 19–21, 2008. (Electronic)

Weber, P. and L. Jouffe. Complex system reliability modeling with Dynamic Object Oriented Bayesian Networks (DOOBN). *Reliability Engineering & System Safety*, **91**:149–162, 2006.

Wilson, A.G., L.A. McNamara, and G.D. Wilson. Information integration for complex systems. *Reliability Engineering & System Safety*, **92**:121–130, 2007.

Chapter 10

Quantifying the benefits of risk reduction in civil infrastructure systems

Erik Vanmarcke

Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA

ABSTRACT: The value and use of probabilistic concepts in civil engineering practice are explored, with a focus on an action-oriented approach in which the engineer identifies damage or failure risks and their consequences, and quantifies the effectiveness and benefits of data-acquisition, design, or monitoring strategies in reducing risk. Applications mentioned relate to various types of civil infrastructure systems, subjected to multiple hazards.

1 Introduction

Uncertainty is a fact of life in the practice of civil engineering. Material properties often differ from those assumed during analysis and design, while loads and environmental conditions clearly defy accurate prediction. Probability, interpreted in the broad sense, can express one's state of knowledge about uncertain events (that may affect the performance and safety of infrastructure systems). At present, the profession most often relies on codes and procedures that not require explicit, quantitative treatment of risk. To compensate for uncertainties, conservative assumptions are usually made about material properties, loads, and analysis methods. As uncertainties are not formally quantified, however, the overall conservatism may vary greatly from project to project, for different types of hazards, and in different decision situations for the same project. Failure to make adequate allowance for uncertainty may jeopardize safety, while the reverse may amount to "wasteful over-conservatism and less satisfactory solutions than if reasonable risks were accepted" (Peck, 1977).

Widely used deterministic formats of design or performance specifications may impede rational risk assessment and communication about risk. Simple code formats consider a design acceptable, or a design provision satisfied, if the computed safety factor exceeds a prescribed or allowable value. Such a criterion labels structures as either safe or unsafe, connoting that all alternative designs that satisfy the criterion are safe. Little thought is given, as a result, to the ever-present residual chance of failure, the factors affecting it, and the opportunities to reduce it. Rigid factor-of-safety provisions, by taking responsibility for the decision-making away from the engineer, serve, in a way, as cookbooks, or worse, as crutches.

Typical multi-criteria codes may be far from optimal, considering how code provisions emerge and evolve. Individual modes or mechanisms of failure, or types of hazards, are often treated separately. (The term "failure", as used herein, is equivalent to "reaching a limit state"; its likelihood, during a specified time interval, is the "failure risk"). Each code provision may be the responsibility of a different professional group

or committee, and code changes tend toward greater conservatism; a provision is likely to be judged satisfactory if failure type (mode, mechanism, hazard) it protects against no longer occurs. If it does, the tendency is to boost conservatism. Designs may thus become over-conservative with respect to certain failure modes (amenable to analysis), and lack balance with respect to the reliability of the "system" as a whole.

1.1 *The sequential nature of decisions*

Consider a typical sequence of decisions encountered, for instance, in a foundation engineering project: (a) Exploration and testing: choose the types of soil tests and the number and spacing of borings; (b) Design: select either a shallow or a deep foundation, and choose its dimensions and materials; (c) Site improvement: decide on whether to compact, preload, or replace questionable materials; (d) Construction: select methods construction and plan compaction control; (e) Surveillance: decide what to measure, where, how often, and for how long, and what to do with the measurements.

Since quantity and quality of information at hand differ much, as do the time and resources available to analyze it, the engineer must, in each decision situation, weigh the cost of alternative actions against their effectiveness in reducing (and better defining) risks. The decisions are closely related, sequential, and, to some extent, hierarchical. For instance, a hazardous foundation condition such as (so-called) solution cavities in limestone areas, is at first only suspected, then clarified during exploration and construction, and may have to be monitored during operation. Assuming records are kept, the amount and quality of the information about site conditions tends to increase with time. Estimates of risk and reliability are, in this sense, time-dependent. Predictions of performance (whether in terms of safety factors, reliability indices, or probabilities) made at the design stage can, in principle, be updated based on evidence obtained during construction and operation; they also depend on time-varying loading and environmental conditions. Probability theory provides a mathematical procedure, Bayes' theorem, for updating "prior" predictions – it parallels, and in a sense formalizes, the "observational approach".

1.2 *Integrated risk assessment methods*

Risk-based decision analysis provides a framework within which the engineer identifies the kinds and degrees of risk involved (at some stage) in a project, and the consequences should "failure" occur, and seeks to evaluate the effectiveness of alternative actions (e.g., in site selection, design, construction, or monitoring) aimed at reducing and controlling risk.

In general, three different methods can be used to do integrated, multi-hazard risk assessment of complex engineered systems:

1.2.1 *Event tree – tracing event sequences "forward"*

This approach is most useful for analyzing the safety and performance of structures subjected to "active" hazards such as natural or accidental (man-made) loads or environmental conditions, e.g., earthquakes, typhoons, floods, and impacts or explosions. The analyst traces event sequences that may lead to failure forward in time. Initiating events with prescribed intensity are usually characterized by annual probabilities of recurrence or "mean return periods".

1.2.2 *Fault tree – tracing event sequences "backward"*

This approach focusing on "system failure" (called the "top event" in fault tree theory); one traces all event sequences that could lead to system failure backward. The approach is capable of dealing with "passive hazards" (e.g., due to material deterioration) and is particularly well suited for use in decision-making about existing structures.

1.2.3 *System reliability analysis*

In structural engineering, this analysis is usually applied to a multi-member structure subjected to a particular random load or combination of loads. The approach relates most directly to conventional design codes in that each code constraint corresponds to a particular "limit state" characterized by its "modal" reliability, not just by a modal safety factor (or partial safety factors). System reliability is concerned with the interaction between the different modes and their local or system-wide consequences. Methods applicable to discrete-member systems can be extended, based on random field theory and stochastic finite elements (Vanmarcke 2009) to reliability analysis of systems in which loads and resistances vary randomly in space as well as with time.

1.3 *Comparison of methods*

The event tree and fault tree approaches may be thought of as "macroscopic" and "action-oriented", as they require relatively little knowledge of mathematical probability and statistics and are suitable for expressing engineering judgment (about probabilities). The fault tree approach, in particular, as is shown below, provides a ready framework for quantifying the benefits of actions aimed primarily at risk reduction; while the costs of hazard mitigation programs are often plain, their benefits – in terms of reduced expected future losses – are not, thus often hindering their adoption.

The system reliability approach, by contrast, is seen as "microscopic" and "analysis-oriented"; it seeks to account explicitly for variability in material properties and loads and predict system performance fully and formally in probabilistic terms.

The methods are of course interrelated, and the risk-related elements in the procedures described in the next section (which are based on fault tree analysis) can in many cases be determined (or better estimated) by using event tree and system reliability tools.

2 Risk-Based decision analysis

2.1 *Elements in the analysis*

2.1.1 *Relative (or fractional) risks*

Consider an existing structure with annual failure risk, p, expressed as a sum of contributions due to each of the causative hazards (or failure modes), $j = 1, 2, \ldots$; we may write

$$p = \sum p_j, \tag{1}$$

where p_j denotes the risk of failure due to hazard j. All possible failure modes must be represented in the summation; this can be achieved by lumping all unknown or

unidentified failure causes into a single category labeled "miscellaneous hazards". The summation implies that mode failure events cannot occur simultaneously; in other words, hazards must be defined so as to be mutually exclusive. [Note: in fault tree theory, equation (1) can also be interpreted in terms of "minimal cut sets".]

Note that the relative (or fractional) risks p_j/p (where $j = 1, 2, \ldots$) sum to one; each value p_j/p is the likelihood that failure, if it happens, will be caused by hazard j. These relative risks can be estimated from data on past failures (and near-failures), using professional judgment and project-specific information (such as inspection reports). Relative risks should be similar for structures of the same type in the same region, with similar site conditions.

For instance, in the case of an existing earth dam, system failure – sudden release of the contents of the reservoir – may be caused by overtopping, piping (or internal erosion), sliding (not earthquake-related), earthquake-induced sliding, and (the above-mentioned category) "miscellaneous hazards", for a total of five hazard categories, $j = 1, 2, \ldots, 5$. These relative risks should be similar for older, non-engineered earthen dams located in a region with, say, moderate seismicity (Vanmarcke, 1974).

2.1.2 *Reference alternative; consequences; risk cost*

In analyzing the benefits of actions aimed at risk reduction, it is useful to express both costs and risks in relation to an existing "as is" (or *status quo*) condition; this reference "do nothing" alternative involves some (often poorly known) annual occurrence probability, p. Let C_m denote the expected monetary loss if failure occurs; the product $C_m p$ is the annual "risk cost", namely the expected annual loss due to (possible) failure. In some cases, it makes sense to assume C_m to be the same regardless of which hazard causes the failure, and that protective actions will affect only p, not C_m. An example would be the case of an existing dam, where the reference action may be to "do nothing", while C_m is the estimated economic loss (at the dam site and downstream) in the event of dam failure.

2.1.3 *Effectiveness of mitigating actions*

As mentioned, engineers generally have little experience with quantifying the benefits of added protection; these take the form of reduced failure risk, and reduced likelihood of losses. Denoting the failure chance *with* and *without* added protection by p^* and p, respectively, we can define the risk reduction effectiveness measure r by means of the relationship

$$p^* = p(1 - r), \tag{2}$$

where r is the fraction of the reference-action risk p that is eliminated by the mitigating action. Hence, $r = 0$ means the action is totally ineffective, as it implies $p^* = p$; there is no change in the risk. The value $r = 1$ indicates 100% effectiveness; the risk is eliminated ($p^* = 0$). An action for which $r = 0.9$ reduces the risk by an order of magnitude. A negative value of r implies $p^* > p$; the action increases the level of risk.

Specific mitigating actions often aim at limiting the risk posed by a specific hazard; for instance, raising the crest of a dam reduces the overtopping risk, while adding a berm aims at reducing the probability of embankment sliding). The implication is that

the effectiveness of mitigating actions is most easily and informatively quantified for specific types of hazards. To capitalize on this, it is useful to express the risk p^*, like p itself, as a sum of contributions due to all the failure causes $j = 1, 2, \ldots$, as follows:

$$p^* = \sum_j p_j^* = \sum_j p_j(1 - r_j), \tag{3}$$

where p_j^* denotes the (estimated) risk due to hazard j if the mitigating action is taken, and r_j is the action's effectiveness in reducing the risk due to hazard j. (By definition, $r = r_j = 0$ for the "do nothing" action). The overall effectiveness index r can then be expressed as a weighted combination of the r_j values and the relative risks p_j/p:

$$r = \sum_j (p_j/p) r_j. \tag{4}$$

To achieve higher overall risk-reduction effectiveness r, available funds should be spent on mitigation of hazards for which fractional risks p_j/p are high, and on actions yielding high corresponding effectiveness indices r_j. Estimating those values (for both p_j/p and r_j) will require a combination of professional experience and judgment, data from historic failures and near-failures (owing to different hazards), and perhaps also support from formal reliability analysis. Specific actions may be highly effective with respect to a "target" hazard i, (so that r_i may be close to one), and ineffective with respect to all the other hazards ($r_j \approx 0$ for $j \neq i$). The "miscellaneous hazards" category, for which any action may be regarded as ineffective, then produces an upper bound, that can be evaluated by means of equation (3), on the overall effectiveness r.

The approach is particularly valuable when negative values of r_j are involved; that is, when a particular action has negative impact on one (or more) of the contributions to the overall risk. For instance, raising the crest of an existing earth dam, which a hydrologist may recommend, reduces the probability of overtopping but increases the risk of sliding; whether this action is justifiable, regardless of cost, will depend on the relative risks of overtopping and sliding, and the values of the corresponding effectiveness indices (r_j). The analysis clearly invites interdisciplinary communication – between hydrologists and geo-engineers, in the case just mentioned – in the interest of balanced hazard mitigation. The common way of pursuing safety, based on discipline-specific guidelines, however well intended, leads to wasteful spending and underachievement in risk reduction.

2.1.4 *Expected benefit*

The average annual monetary losses *with* and *without* added protection are $C_m p^*$ and $C_m p$, respectively, and their difference is the annual average economic benefit of the risk reduction:

$$b = C_m p - C_m p^* = C_m p r. \tag{5}$$

In words, b is the product of the *status quo* risk p, the hazard potential C_m, and the action's overall effectiveness index r.

2.1.5 *Added cost for mitigation*

The last element in the analysis – the one probably most familiar to decision makers – is the cost of providing added protection. Each action is characterized by an added-cost-per-year Δc. (Given a discount rate and a time horizon, capital expenditures can also be expressed as an annual disbursement Δc, like a "mortgage payment".)

2.2 *Basic format of the analysis*

The effectiveness index r and the annualized added cost Δc (or some equivalent cost index) must be evaluated for each action. One can construct a simple (fractional effectiveness) *matrix* whose elements are the measures of fractional effectiveness r_{jk} of each mitigating action k in reducing the risk due to hazard j. The matrix has one column for each hazard ($j = 1, 2, \ldots$); these are listed on top of each column, along with the corresponding relative risk (p_j/p) – these must of course sum to one: $\sum_j (p_j/p) = 1$.

For each (new) mitigating action, one row is added to the matrix. Using equation (4), one obtains for each alternative k the overall effectiveness $r \equiv r_k$. The best action, on an expected cost basis, is the one maximizing the expected net benefit $[b - \Delta c]_k$. Also, any action for which $[b - \Delta c]_k$ is positive, is preferable to the "do nothing" alternative. It is also useful to plot the quantities b_k, Δc_k, and $[b - \Delta c]_k$ (for each action k) against the action's overall effectiveness r_k, taking note of the value of r_k that maximizes $[b - \Delta c]_{\mathrm{k}}$.

The "do nothing" alternative costs nothing initially ($\Delta c = 0$) but also brings no (annual expected) benefit ($b = 0$, since $r = 0$); it may have high annual future expected cost ($C_m p^* = C_m p$). On the opposite side of the effectiveness scale, achieving a value such as $r = 0.99$ may be prohibitively expensive, or, in case there are "unknown cause" risks for which $p_j/p > 0.01$, judged impossible.

If, by applying equation (4), an action's overall effectiveness measure r comes out negative, then that action should obviously not be taken, regardless of cost. It is easy to see how that could happen, say, in a case involving action aimed at reducing the risk of failure of an existing levee. Heightening the levee's crest (without also adding a berm) protects against inundation ($r_1 > 0$) but worsens the levee's stability ($r_2 < 0$). If the two relative risks are judged to equal, $p_1/p = p_2/p$, and there happen to be no other significant contributions to the overall risk of failure, then the overall effectiveness r will be negative in case $r_1 < -r_2$.

2.3 *Extensions of the basic format*

2.3.1 *"Risk cost" format*

In general, failure consequences may depend on the (type of) hazard j and be changed by the mitigating actions taken. Instead of doing the analysis in terms of probabilities p and p_j, it is then more productive to replace the latter by the "risk costs" $q_j \equiv p_j C_{m,j}$, the risk cost for hazard j, and the sum (over all j) of all these by q. The risk-cost ratios, q_j/q, whose sum is one, now play the same role as the relative risks (and are identical to them if $C_{m,j} = C_m$ for all j). All indices of effectiveness now reflect the impact on risk costs, but the analysis format remains the same: the p's are all replaced by q's in equations (1) through (4).

2.3.2 *Non-monetary consequences*

The methodology can also be used to quantify benefits of hazard mitigation measures in terms of lives saved (or injuries prevented); the monetary consequences of failure, C_m, are replaced by the life loss potential C_l (expected number of fatalities if failure occurs). A warning system at a dam or levee site, for instance, will be aimed mainly at preventing life loss (by providing timely warning); hence its effectiveness in reducing life loss may be close to one, while its effectiveness in reducing property loss is close to zero.

2.3.3 *Use in making design decisions*

The methodology, in the format presented, applies not only to existing structures; in a design situation, an appropriate reference alternative may be a "standard design" or a "preliminary" or "trial" design. One now considers design changes, actions that differ from the reference alternative, and evaluates their impact on the (initial) cost and on the risk and consequences of failure.

2.3.4 *Optimizing risk reduction programs involving a system of existing structures*

The methodology can be used to deal with situations of great practical interest, where risk-informed decisions need to be made about inspection, maintenance and repair of a group of structures (such as offshore structures, bridges, or dams) subjected to multiple hazards. We now use the subscript i to refer to a specific structure in the group. The total expected annual economic benefits of a safety program can be evaluated by summing the benefits b_i associated with each structure. (The probable total annual number of lives saved can be evaluated similarly).

For example, in the case of offshore structures, it may be reasonable to adopt a (default) value for the average annual risk of failure of a typical offshore structure. (More generally, based on examination of available failure statistics and engineering judgment, the analyst may wish to refine such an estimate by allowing it to depend on structural type, age, and design criteria). For each category of structures for which a set of relative risks is developed, one should attempt to construct a matrix of values the effectiveness index r_{ij} for every alternative monitoring (or repair) strategy considered for that structural category. These values, typically between 0 and 1, indicate the fractional amounts by which the analyst expects risks to be reduced following the implementation of the risk reduction measure. Most r_{ij} values will be close to zero (implying the procedure has little impact on the risk) or close to one (implying a risk reduction by an order of magnitude or higher). They must be determined based on a combination of professional judgment, probability analysis and, to the extent data are available, performance/failure records.

Given a fixed annual budget for a safety program covering many existing structures of different types and sizes, a reasonable objective in designing the program, i.e., in choosing the mix of protective actions to be taken, is to maximize the total expected monetary benefits, assuming economic issues dominate the decision. (If this is not the case, multi-objective decision analysis, in which all hazard-potential components are considered, may be attempted). Examples of the use of the methodology, and related further details, may be found in Erickson et al. (1989); for dams in Vanmarcke and

Bohnenblust (1982) and Bohnenblust and Vanmarcke (1982); and for bridges in Cesare et al. (1993, 1994).

3 Conclusions

The proposed approach to risk-informed engineering decision-making attempts to put technical and socio-economic issues into proper focus by organizing factual information about risks, costs and losses, both monetary and non-monetary. It has the following notable features:

(a) It provides a format for summarizing and transmitting information about past failures, about the relative frequency of various causes of failure, and about the effectiveness of different methods for reducing the risk of failure.
(b) It provides a framework for arriving at balanced decisions in which no single contributing hazard or cost component unduly dominates the mitigation efforts.
(c) It can serve to clarify liability issues: which party (engineer, owner, contractor, the public) is exposed to what type of risk during which phase of a project? Explicit recognition of where the risks originate, how large they are, and how they can be controlled or reduced, should lead to better contracting through agreed-upon assignment of risks in contracts.
(d) It facilitates communication about risk and the cost of risk reduction among the parties involved in decisions about hazard mitigation, including engineers, owners, regulatory agencies, insurance buyers and providers, and the public.
(e) It enables quantification (in monetary terms) of the benefits of programs and measures aimed primarily at risk reduction (such as health monitoring programs).

In this broad, varied context, the concepts and tools of quantitative risk analysis appear essential to the advancement of the art and practice of civil engineering.

References

Bohnenblust, H. and Vanmarcke, E., "Decision Analysis for Prioritizing Dams for Remedial Measures: A Case Study", M.I.T. Dept. of Civil Engineering Research Report R82-12, 1982.

Cesare, M.A., Santamarina, C., Turkstra, C. and Vanmarcke, E., "Risk-Based Bridge Management", *Journal of Transportation Engineering*, ASCE, Vol. 119, No. 5, pp. 742–750, 1993.

Cesare, M.A., Santamarina, C., Turkstra, C. and Vanmarcke, E., "Risk-Based Bridge Management: Optimization and Inspection Scheduling", *Canadian Journal of Civil Engineering*, Vol. 21, pp. 897–902, 1994.

Erickson, H.L., Mulvey, J.M., and Vanmarcke, "Integrating Expert Systems and Mathematical Programming: An Example in Infrastructure Management", *Annals of Operations Research*, Vol. 21, pp. 275–300, 1989.

Peck, R., Pitfalls of Overconservatism in Geotechnical Engineering, *Civil Engineering*, ASCE, pp. 62–66, 1977.

Vanmarcke, E., Decision Analysis in Dam Safety Monitoring, *Proc. Eng. Found. Conf. on the Safety of Dams*, Henniker, New Hampshire, ASCE, pp. 127–148, 1974.

Vanmarcke, E. and Bohnenblust, H., "Risk-Based Decision Analysis in Dam Safety", M.I.T. Dept. of Civil Engineering Research Report R82-11, 1982.

Vanmarcke, E., *Random Fields*, 2nd Ed., World Scientific Publishing, 2009.

Chapter 11

A risk-based framework for multi-hazard management of infrastructure

Bilal M. Ayyub

Center for Technology and Systems Management, Department of Civil and Environmental Engineering, University of Maryland College Park, MD, USA

ABSTRACT: A quantitative risk analysis framework for infrastructure management is introduced to address multi-hazards. The underlying methodology is intended to assist decision and policy makers, and has the characteristics of being analytic, quantitative and probabilistic. A hazard is quantified using a probabilistic framework to obtain hazard intensity exceedance rates, such as flood elevation exceedance rate, and the risk is quantified in the form of loss exceedance rates that are based on a spectrum of hazards. Two cases studies are used to illustrate the framework: a protection system of hurricane-prone region, and security protection of infrastructure. Hurricanes are generated using a joint probability distribution of the parameters that define hurricane intensity, and human-caused hazards are selected to demonstrate the methodology. The proposed methodology will enable decision makers to evaluate alternatives for managing risk, such as providing increased hurricane protection, increasing evacuation effectiveness, changing land-use policy, enhancing hurricane protection system operations, enhancing preparedness for the case of natural hazards, providing secure perimeter, reducing vulnerabilities, reducing threats, and/or enhancing consequence mitigation to simultaneously address multiple hazards.

1 Risk definition and quantification

Within the context of infrastructure, risk is commonly associated with an uncertain event or condition that, if it occurs, has a negative impact on the performance of infrastructure and could lead to adverse consequences.

Formally, risk can be defined as the potential of losses resulting from an exposure to a hazard or as a result of an uncertain event (Ayyub 2003). Risk should be based on identified risk events or event scenarios. Risk can be viewed to be a multi-dimensional quantity that includes event-occurrence probability, event-occurrence consequences, consequence significance, and the population at risk; however, it is commonly measured as a pair of the probability of occurrence of an event, and the outcomes or consequences associated with the event's occurrence. This pairing can be represented by the following equation:

$$Risk \equiv [(p_1, c_1), (p_2, c_2), \ldots, (p_n, c_n)] \tag{1}$$

where p_i is the occurrence probability of an outcome or event i out of n possible events, and c_i is the occurrence consequences or outcomes of the event. A generalized definition of risk can be expressed as

$$Risk \equiv \begin{bmatrix} (l_1, o_1, u_1, cs_1, po_1), \\ (l_2, o_2, u_2, cs_2, po_2), \ldots, \\ (l_n, o_n, u_n, cs_n, po_n) \end{bmatrix} \tag{2}$$

where l is likelihood, o is outcome, u is utility (or significance), cs is causal scenario, po is population affected by the outcome, and n is the number of outcomes. The definition according to Eq. 2 covers all attributes measured in risk assessment, and offers a more complete description of risk compared to Eq. 1, from the cause event to the affected population and consequences. The population-size effect should be considered in risk studies since society responds differently for risks associated with a large population in comparison to a small population. For example, a fatality rate of 1 in 100,000 per event for an affected population of 10 results in an expected fatality of 10^{-4} per event whereas the same fatality rate per event for an affected population of 10,000,000 results in an expected fatality of 100 per event. Although, the impact of the two scenarios might be the same on the society (same risk value), the total number of fatalities per event/accident is a factor in risk acceptance. Plane travel may be *safer* than for example recreational boating, but 200 to 300 injuries per plane accident are less acceptable to society. Therefore, the size of the population at risk and the number of fatalities per event should be considered as factors in setting acceptable risk levels.

Risk is quantified as the rate (measured in events per unit time, such as a year) that lives, economic, environmental, and social/cultural losses will occur due to the non-performance of an engineered system or component. The non-performance of the system or component can be quantified as the probability that specific loads (or demands) exceed respective strengths (or capacities) causing the system or component to fail, and losses are defined as the adverse impacts of that failure if it occurs. Another common representation of risk is in the form of an exceedance rate (or probability) function of consequences. In a simplified notional (or Cartesian) product, risk is commonly expressed as:

$$\text{Risk} = \text{Event rate} \times \text{Vulnerability} \times \text{Consequences} \tag{3}$$

This equation not only defines risk but also offers strategies to control or manage risk: by making the system more reliable through vulnerability reduction or by reducing the potential losses resulting from a failure. The conditional probability of failure in the vulnerability component of the equation can be influenced by engineers by strengthening of existing structures or by adding additional protection; however the consequence part is highly dependent upon the actions and decisions made by residents, government and local officials, including first-response and evacuation plans and practices. For example, in densely populated areas, simply increasing the reliability of a hurricane protection system may not reduce risks to acceptable levels and may increase consequences through continued flood plain development that can offset any risk reductions.

2 Risk analysis framework

Risk studies require the use of analytical methods at the system level that considers subsystems and components in assessing their failure probabilities and consequences. Systematic, quantitative approaches for assessing the failure probabilities and consequences of engineering systems are used for this purpose. A systematic approach allows an analyst to evaluate expediently and easily complex systems under different operational and extreme conditions. The ability to quantitatively evaluate these systems helps cut the cost of unnecessary and often expensive redesign, repair, strengthening or replacement of components, subsystems and systems. The results of risk analysis can also be utilized in decision analysis methods that are based on cost-benefit tradeoffs.

Risk assessment is a technical and scientific process by which the risks of a given situation for a system are modeled and quantified. Risk assessment can require and/or provide both qualitative and quantitative data to decision makers for use in risk management.

Risk assessment or risk analysis provides the process for identifying hazards, event-probability assessment, and consequence assessment. The risk assessment process answers three basic questions: (1) what can go wrong? (2) What is the likelihood that it will go wrong? (3) What are the consequences if it does go wrong? Answering these questions requires the utilization of various risk methods as discussed by Ayyub (2003). The risk management answers the question of (1) what can be done to reduce or control risk, i.e., what are the solutions, and (2) what are the impacts of any proposed solutions on the risk profile of the system.

A risk assessment process should utilize experiences gathered from project personnel including managers, other similar projects and data sources, previous risk assessment models, experiences from other industries and experts, in conjunction with analysis and damage evaluation and prediction tools. A risk assessment process should be employed part of a risk-based or risk-informed methodology constructed as a synergistic combination of decision models, advanced probabilistic reliability analysis algorithms, failure consequence assessment methods, and conventional performance assessment measures. The methodology should realistically account for the various sources and types of uncertainty involved in the decision-making process (Ayyub and McCuen 2003; Ayyub and Klir 2006).

In this section, a typical overall methodology is provided in the form of a workflow or block diagram. The various components of the methodology are described in subsequent sections. Figure 1 provides an overall description of a methodology for risk-based management of structural systems for the purpose of demonstration. The methodology consists of the following primary steps:

- Definition of analysis objectives and systems;
- Hazard analysis, definition of failure scenarios, and hazardous sources and their terms;
- Collection of data in a lifecycle framework;
- Qualitative risk assessment if needed;
- Quantitative risk assessment; and
- Management of system integrity through failure prevention and consequence mitigation using risk-based decision making.

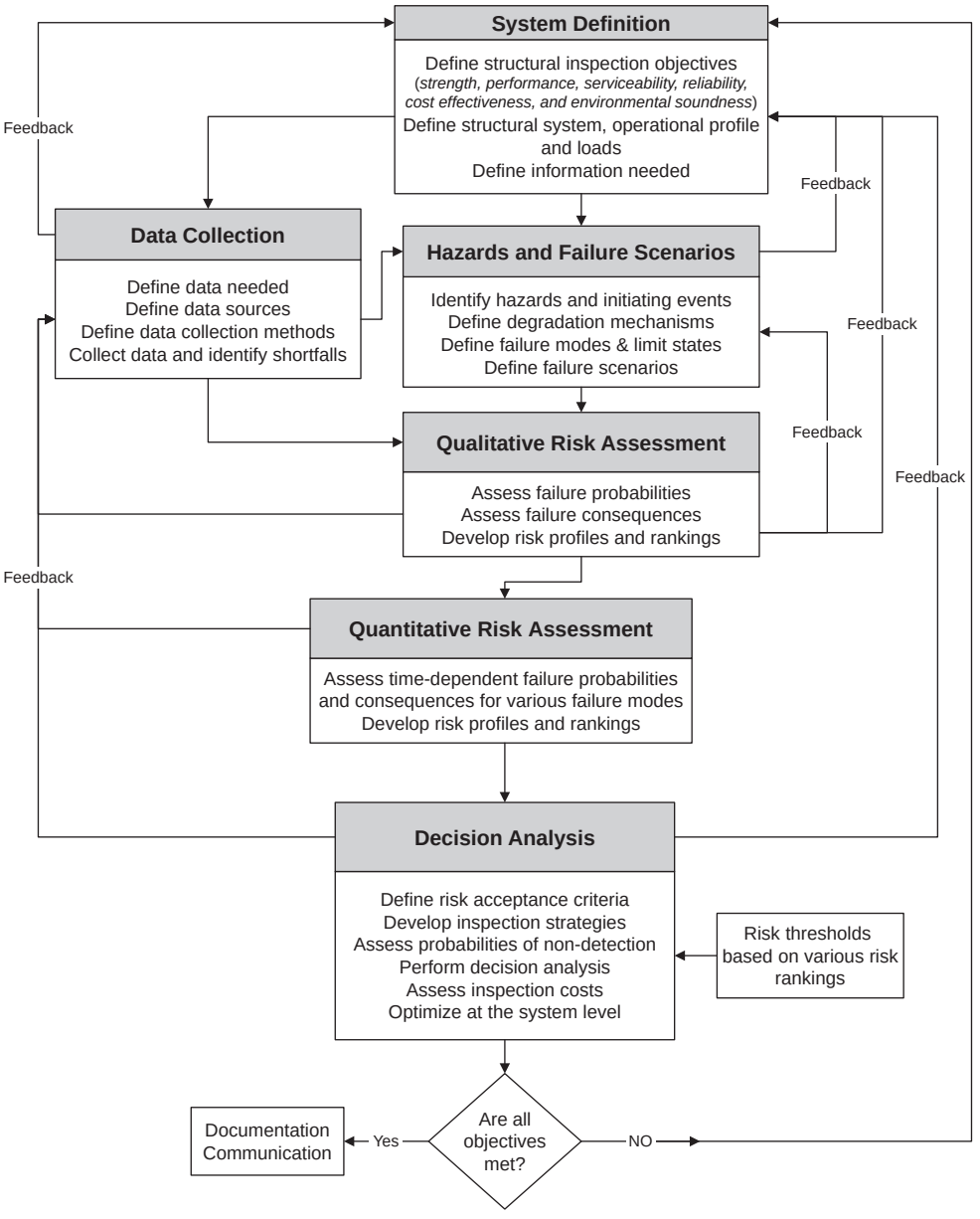


Figure 1 Methodology for risk-based lifecycle management of structural systems.

The first step of the methodology is to define the system. This definition should be based on a goal that is broken down into a set of analysis objectives. A system can be defined as an assemblage or combination of elements of various levels and/or details that act together for a specific purpose. Defining the system provides the risk-based

methodology with the information it needs to achieve the analysis objectives. The system definition phase of the methodology has four main activities. The activities are to

- Define the goal and objectives,
- Define the system boundaries,
- Define the success criteria in terms of measurable performances,
- Collect information for assessing failure likelihood, and
- Collect information for assessing failure consequences.

For example, structural systems require a structural integrity goal that can include objectives stated in terms of strength, performance, serviceability, reliability, cost effectiveness, and environmental soundness. The objectives can be broken down further to include other structural integrity attributes, such as alignment and watertightness in case of marine vessels. A system can be defined based on a stated set of objectives. The same system can be defined differently depending on these stated objectives. A marine vessel structural system can be considered to contain individual structural elements such as plates, stiffened panels, stiffeners, longitudinals, . . . , etc. These elements could be further separated into individual components and/or details. Identifying all of the elements, components and details allows an analysis team to collect the necessary operational, maintenance and repair information throughout lifecycle on each item so that failure rates, repair frequencies and failure consequences can be estimated. The system definition might need to include non-structural subsystems and components that would be affected in case of failure. The subsystems and components are needed to assess the consequences.

In order to understand failure and the consequences of failure, the states of success need to be defined. For the system to be successful, it must be able to perform its designed functions by meeting measurable performance requirements. But the system may be capable of various levels of performance, all of which might not be considered a successful performance. While a marine vessel may be able to get from point A to point B only at a reduced speed due to a fatigue failure that results in excessive vibration at the engine room, its performance would probably not be considered successful. The same concept can be applied to individual elements, components and details. It is clear from this example that the vessel's success and failure impacts should be based on the overall vessel performance that can easily extend beyond the structural systems.

With the development of the definition of success, one can begin to assess the likelihood of occurrence and causes of failures. Most of the information required to develop an estimate of the likelihood of failure might exist in maintenance and operating histories available on the systems and equipment, and based on judgment and expert opinion. This information might not be readily accessible, and its extraction from available sources might be difficult. Also, assembling it in a manner that is suitable for the risk-based methodology might be a challenge.

Operation, maintenance, engineering and corporate information on failure history should be collected and analyzed for the purpose of assessing the consequences of failures. The consequence information might not be available from the same sources as the information on the failure itself. Typically there are documentations of repair costs, re-inspection or re-certification costs, lost person-hours of labor, and possibly even lost opportunity costs due to system failure. Much more difficult to find and assess

are costs associated with the effects on other systems, the cost of shifting resources to cover lost production, and things like environmental, safety-loss or public-relations costs. These may be attained through carefully organized discussions and interviews with cognizant personnel including the use of expert-opinion elicitation.

Risk assessment methods can be categorized according to how the risk is determined into quantitative or qualitative analysis. Qualitative risk analysis uses judgment and sometimes "expert" opinion to evaluate the probability and consequence values. This subjective approach may be sufficient to assess the risk of a system, depending on the available resources.

Quantitative analysis relies on probabilistic and statistical methods, and databases that identify numerical probability values and consequence values for risk assessment. This objective approach examines the system in greater detail to assess risks.

The selection of a quantitative or qualitative method depends upon the availability of data for evaluating the hazard and the level of analysis needed to make a confident decision. Qualitative methods offer analyses without detailed information, but the intuitive and subjective processes may result in differences in outcomes by those who use them. Quantitative analysis generally provides a more uniform understanding among different individuals, but requires quality data for accurate results. A combination of both qualitative and quantitative analyses can be used depending on the situation.

Risk assessment requires estimates of the vulnerabilities and failure likelihood at some identified levels of decision-making. The failure likelihood can be estimated in the form of lifetime failure likelihood, annual failure likelihood, mean time between failures, or failure rate. The estimates can be in numeric or non-numeric form. An example numeric form for an annual failure probability is 0.00015, and for a mean time between failures is 10 years. An example non-numeric form for "annual failure likelihood" is large, and for a "mean time between failures" is medium. In the latter non-numeric form, guidance needs to be provided regarding the meaning of terms such as large, medium, small, very large, very small, etc. The selection of the form should be based on the availability of information, the ability of the personnel providing the needed information to express it in one form or another, and the importance of having numeric versus non-numeric information in formulating the final decisions.

Vulnerabilities defined notionally as a weakness in the system for a particular hazard can be estimated in the form of fragilities, i.e., conditional probabilities as functions of hazard or load levels to the system. In the case of security risk, vulnerabilities should also include the probability that a human-caused hazard would succeed, i.e., probability of success of an attack.

The types of failure consequences that should be considered in a study need to be selected. They can include production loss, property damage, environmental damage, and safety loss in the form of human injury and death. Consequences at the identified levels of decision making should be estimated. The estimates can be in numeric or non-numeric form. An example numeric form for production loss is 1000 units. An example non-numeric form for production loss is large. In the latter non-numeric form, guidance needs to be provided regarding the meaning of terms such as large, medium, small, very large, very small, etc. The selection of the form should be based on the availability of information, the ability of the personnel providing the needed information to express it in one form or another, and the importance of having numeric versus non-numeric information in formulating the final decisions.

Table 1 Methods for determining risk acceptance.

Risk acceptance method	*Summary*
Risk Conversion Factors	This method addresses the attitudes of the public about risk through comparisons of risk categories.
Farmers Curve	It provides an estimated curve for cumulative probability risk profile for certain consequences (e.g., deaths).
Revealed Preferences	Through comparisons of risk and benefit for different activities, this method categorizes society preferences for voluntary and involuntary exposure to risk.
Evaluation of Magnitude of Consequences	This technique compares the probability of risks to the consequence magnitude for different industries to determine acceptable risk levels based on consequence.
Risk Effectiveness	It provides a ratio for the comparison of cost to the magnitude of risk reduction.
Risk Comparison	This risk acceptance method provides a comparison between various activities.

3 Risk control and management

Adding risk control to risk assessment produces risk management. Risk management is the process by which system operators, managers, and owners make safety decisions, regulatory changes, and choose different system configurations based on the data generated by risk assessment. Risk management involves using information from risk assessment to make educated decisions about system safety. Risk control includes failure prevention and consequence mitigation.

Risk management addresses the optimal allocation of available resources in support of a goal, and may require the definition of acceptable risk, and comparative evaluation of options and/or alternatives for decision making. The goal of risk management is to reduce risk to an acceptable level and/or prioritize resources based on comparative analysis. Risk reduction is accomplished by preventing an unfavorable scenario, reducing the frequency, and/or reducing the consequence. Strategy tables are commonly used to develop alternatives.

Risk acceptance constitutes a definition of safety, and its complexity in terms of definition and the process to define it could make it controversial. If a system has a risk value above the risk acceptance level, actions should be taken to address safety concerns and improve the system through risk reduction measures. One difficulty with this process is defining acceptable safety levels for activities, industries, structures, etc. Since the acceptance of risk depends upon society perceptions of risks and its values, the acceptance criteria do not depend on the risk value alone. Table 1 summarizes several methods that have been developed to assist in determining acceptable risk values.

According to the "Evaluation of Magnitude of Consequences" technique, a key factor affecting the acceptance of risk is the magnitude of consequence of the event that can result from failure. This technique has been used in several industries to demonstrate the location of the industry within societies' risk acceptance levels based on consequence magnitude as shown in Figure 2. Further evaluation has resulted in several

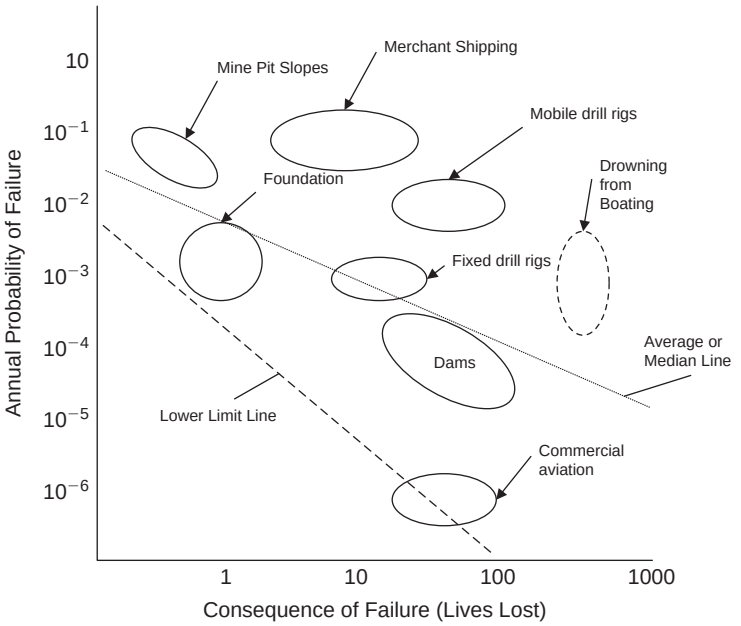


Figure 2 Risks based on consequence of failure for various industries.

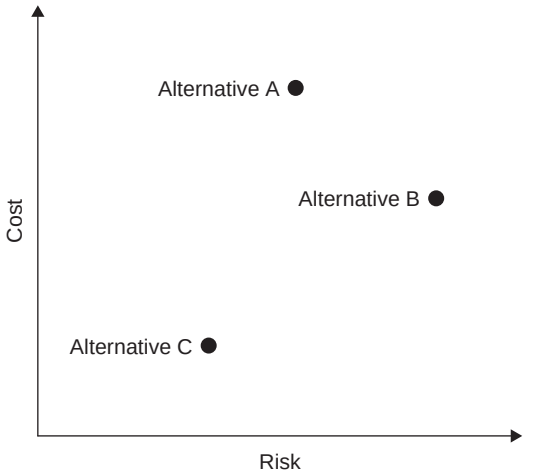


Figure 3 Risks and costs.

estimates for the relationship between the accepted probability of failure and the magnitude of consequence for failure. For example, Figure 2 shows an average or median line and a minimum line. These lines are schematically sketched for illustration purposes.

Risk managers commonly weigh various factors including cost and risk. An analysis of three different alternatives is shown graphically in Figure 3 as an example. The

graph shows that alternative (C) is the best choice since the level of risk and cost is less than alternatives (A) and (B). However, if the only alternatives were A and B, the decision would be more difficult. Alternative (A) has higher cost and lower risk than alternative (B); alternative (B) has higher risk but lower cost than alternative (A). A risk manager needs to weigh the importance of risk and cost in making this decision and availability of resources, and make use of risk-based decision analysis.

The benefit of a risk mitigation alternative or action can be assessed as follows:

$$\text{Benefit} = \text{unmitigated risk} - \text{mitigated risk} \tag{4}$$

The cost in Eq. 4 is the cost of the mitigation action. The benefit minus the cost of mitigation can be used to justify the allocation of resources. The benefit-to-cost ratio can be computed, and may also be helpful in decision-making. The benefit-to-cost ratio can be computed as

$$\frac{\text{Benefit}}{\text{Cost}} = \frac{\text{Unmitigated Risk} - \text{Mitigated Risk}}{\text{Cost of Mitigation Action}} \tag{5}$$

Ratios greater than one are desirable. In general, the larger the ratio, the better the mitigation action. Other decision consideration could include the internal rate of return, initial cost, and political, legal and budgetary considerations.

The previous model for benefit-cost analysis does not account for the full probabilistic characteristics of the benefit (B) and cost (C). Assuming B and C to normally distributed, a benefit-cost index ($\beta_{B/C}$) can be defined as follows (Ayyub 2003):

$$\beta_{B/C} = \frac{\mu_B - \mu_C}{\sqrt{\sigma_B^2 + \sigma_C^2}} \tag{6}$$

where μ and σ are the mean and standard deviation. The failure probability can be computed as

$$P_{f,B/C} = P(C > B) = 1 - \Phi(\beta) \tag{7}$$

In the case of lognormally distributed B and C, the benefit-cost index ($\beta_{B/C}$) can be computed as

$$\beta_{B/C} = \frac{\ln\left(\frac{\mu_B}{\mu_C}\sqrt{\frac{\delta_C^2 + 1}{\delta_B^2 + 1}}\right)}{\sqrt{\ln\left[(\delta_B^2 + 1)(\delta_C^2 + 1)\right]}} \tag{8}$$

where δ is the coefficient of variation. Equation 7 also holds for the case of lognormally distributed B and C. In the case of mixed distributions or cases involving basic random variables of B and C, the advanced second moment method or simulation method can be used (Ayyub 2003). In cases where benefit is correlated with cost, other methods can be used (Ayyub 2003).

4 Risk communication

Risk communication can be defined as an interactive process of exchange of information and opinion among stakeholders such as individuals, groups, and institutions. It often involves multiple messages about the nature of risk or expressing concerns, opinions, or reactions to risk managers or to legal and institutional arrangements for risk management. Risk communication greatly affects risk acceptance.

The process of risk communication can be enhanced and improved in three aspects: (1) the process, (2) the message, and (3) the audiences. The risk assessment and management process needs to have a clear goal with openness, balance, and competence. The content of the message should account for audience orientation and uncertainty, provide risk comparison, and be complete. There is a need to guide and introduce risks associated with a specific technology, the process of risk assessment and management, acceptable risk, decision making, uncertainty, costs and benefits, and feedback mechanisms. Improving risk literacy is an essential component of the risk communication process. The following are guiding considerations in communicating risk:

- Risk communication must be free of jargon,
- Consensus of expert needs to be established,
- Materials cited, and their sources must be credible,
- Materials must be tailored to audience,
- The information must be personalized to the extent possible,
- Motivation discussion should stress a positive approach and the likelihood of success, and
- Risk data must be presented in a meaningful manner.

5 Examples

5.1 *Protecting a hurricane-prone region*

Risk associated with a hurricane protection system (HPS) is quantified through a regional hurricane rate (λ) and the probability $P(C > c)$ with which a consequence measure (C) exceeds particular levels (c). The loss-exceedance probability per event is evaluated as:

$$P(C > c) = \sum_i \sum_j P(h_i)P(S_j|h_i)P(C > c|h_i, S_j) \tag{9}$$

An annual loss-exceedance rate can be estimated as follows:

$$\lambda(C > c) = \sum_i \sum_j \lambda P(h_i)P(S_j|h_i) \times P(C > c|h_i, S_j) \tag{10}$$

where $P(h_i)$ is the probability of hurricane events of type i, $P(S_j|h_i)$ is the probability that the system is left in state j from the occurrence of h_i, and $P(C > c|h_i, S_j)$ is the probability that the consequence C exceeds c under (h_i, S_j). Summation is over all hurricane types i and all system states j in a suitable discretization. Simulation studies of hurricanes for risk analysis require the use of representative combinations of hurricane parameters

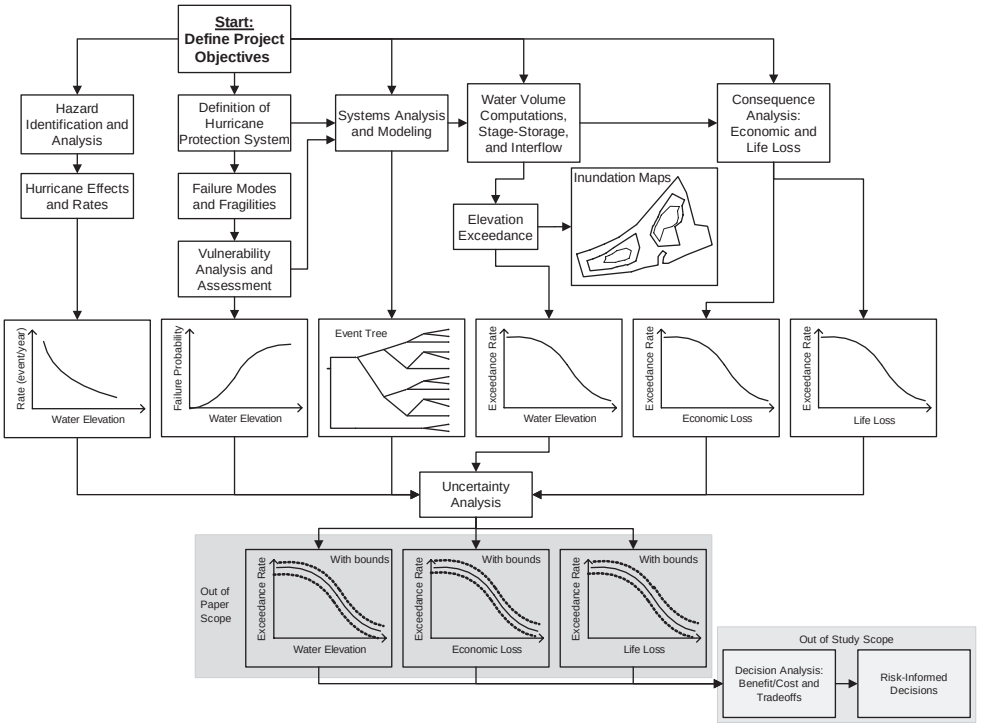


Figure 4 A methodology for hurricane protection systems.

and their respective probabilities. The outcome of this process is a set of hurricane simulation cases and their respective conditional rates $\lambda P(h_i)$.

Evaluation of the regional hurricane rate λ and the probability $P(h_i)$, the conditional probabilities $P(S_j|h_i)$, and the conditional probabilities $P(C > c|h_i, S_j)$ is the main objective of the hurricane model, the system model, and the consequence model, respectively. The probability $P(S_j|h_i)$ covers the states of the components of the HPS, such as closure structure and operations, precipitation levels, electric power availability, failures modes of levees and floodwalls, and pumping station reliability. To assess the state of the HPS given a hurricane event requires an evaluation of the reliability of individual structures, systems and components (e.g., levees, floodwalls, pump systems) when they are exposed to the loads and effects of the hurricane (e.g., the peak surge, wave action) and the relationship of these elements to the overall function of the system to prevent flooding in protected areas.

The methodology and logic tree of Figures 4 and 5 were constructed to determine the rate of flooding elevations and displaying the results as inundation contours within the basins. The processes of transforming inundation to consequences is simplified by grouping communication, warning decision and public execution into an exposure factor parameter applied to lives and property at risk, and grouping power and pumping availability into one event. The events of the tree are defined in Table 2.

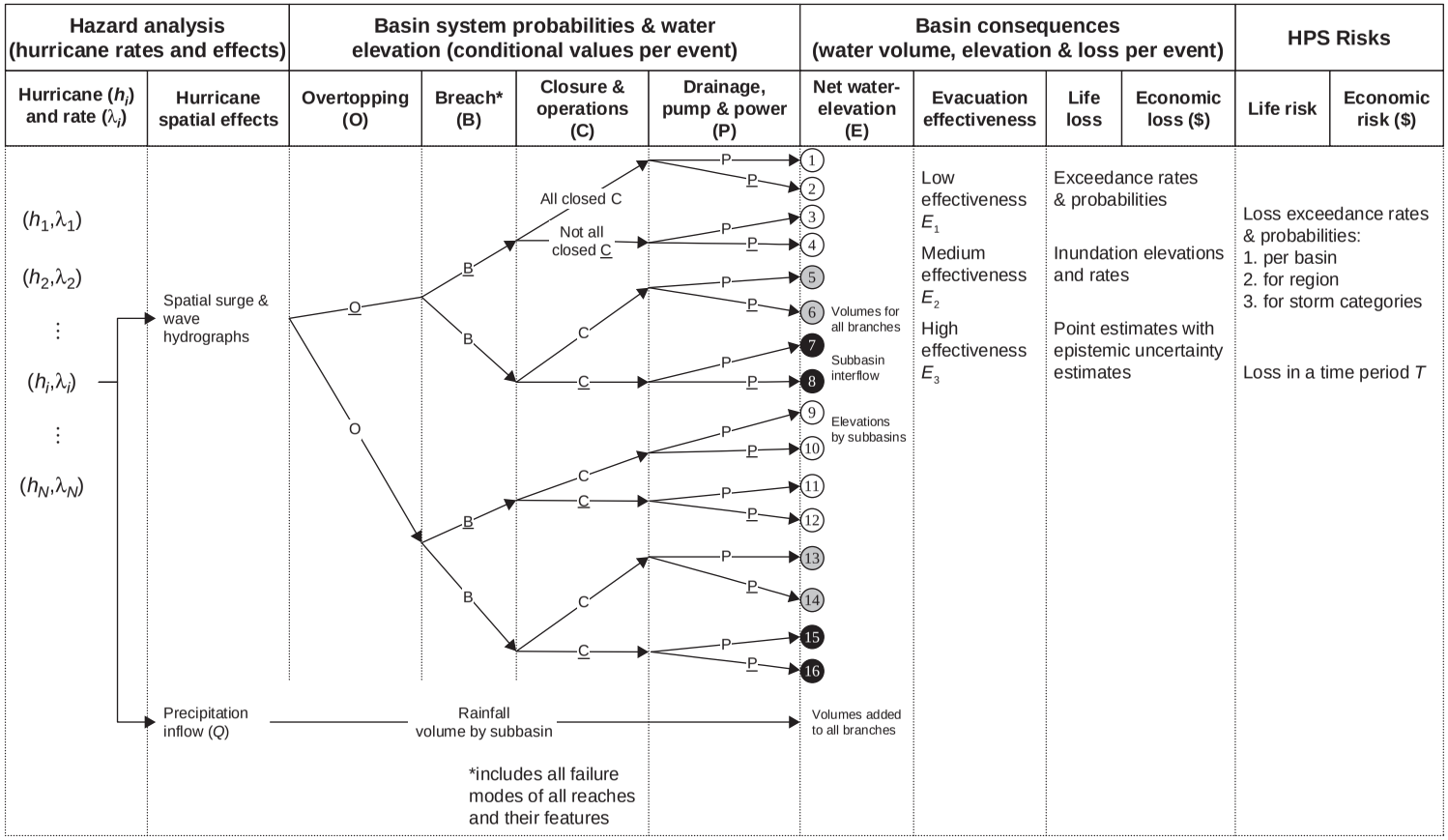


Figure 5 A logic tree for a protected hurricane-prone region.

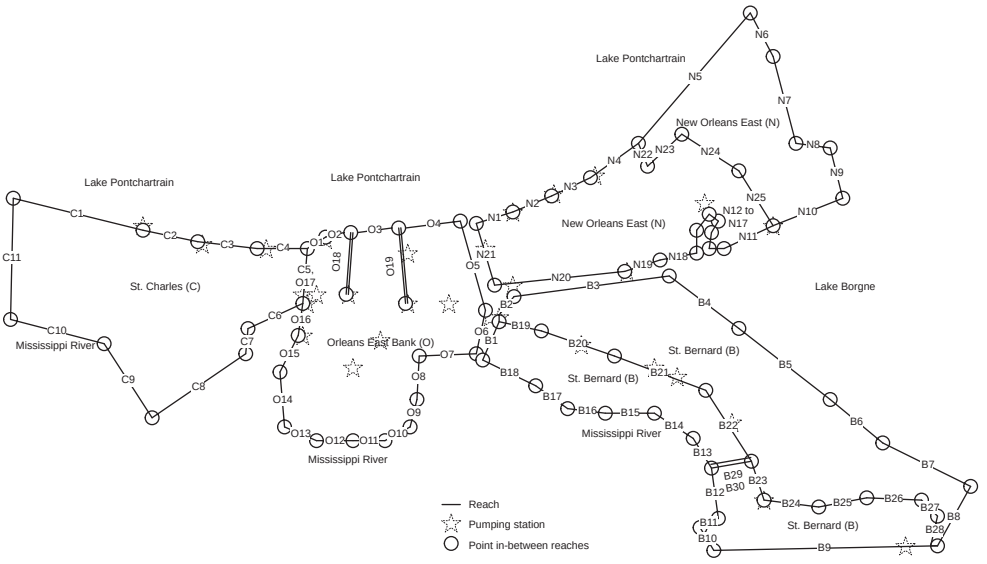


Figure 6 System definition.

Table 2 Tree top events.

Top event	*Description*
Hurricane initiating event	The hurricane initiating event is mapping of hydrographs of the peak flood surge with waves in the study area with a hurricane rate λ. This event was denoted, $h_i(x, y)$, and has a probability of occurrence, $P(h_i(x, y))$ and a rate of occurrence of $\lambda P(h_i(x, y))$.
Closure structure and operations (C)	This event models whether the hurricane protection system closures, i.e., gates, have been sealed prior to the hurricane.
Precipitation inflow (Q)	This event corresponds to the rainfall that occurs during a hurricane event. The precipitation inflow per subbasin is treated as a random variable.
Drainage, pumping and power (P)	This event models the availability of power (normal) power for the pump systems.
Overtopping (O)	This event models the failure of the enclosure/protection system due to overtopping, given that failure has not occurred by some other (non-overtopping) failure mode. If failure (breach) does not occur, flooding due to overtopping could still result.
Breach (B)	This event models the failure of the enclosure/protection system (e.g., levees/floodwalls, closures) during the hurricane, exclusive of overtopping failures).

The hurricane protection system was defined in terms of reaches of floodwalls and levees, basins and subbasins as illustrated in Figure 6 for a hypothetical region. Also, the system definition could include transitions, gates, pumping stations, and other features. Surge and wave hydrographs were simulated using a set of storms that are

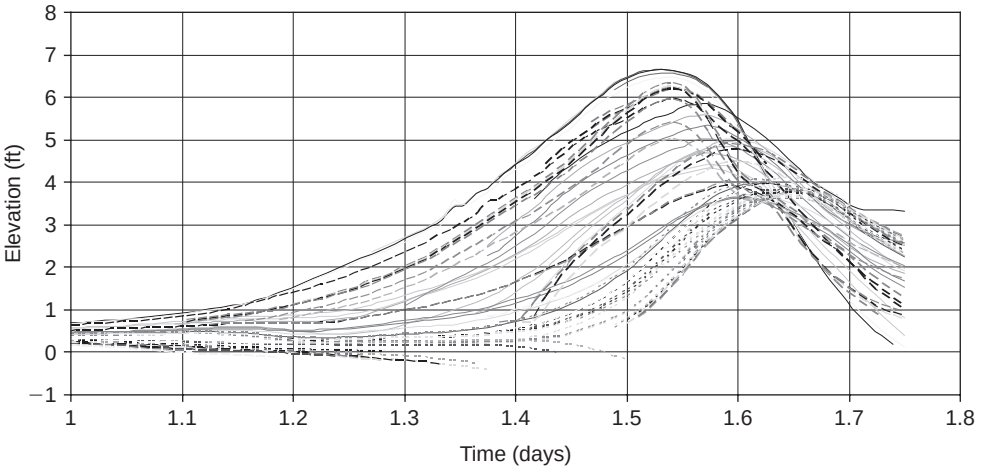


Figure 7 Hydrographs.

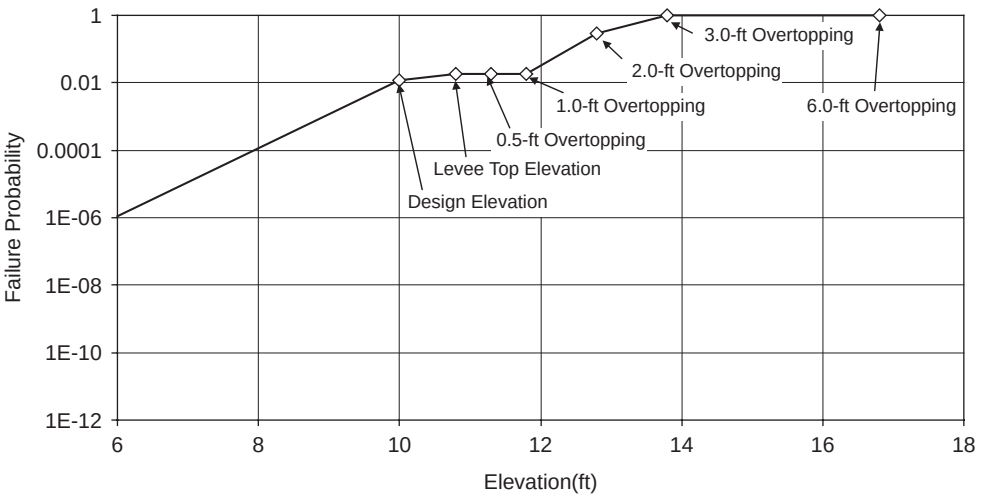


Figure 8 Fragility curves.

representative of all storms with associated rates of occurrence. Example surge and wave hydrographs are provided in Figure 7.

Fragility curves were used for all the reaches and transitions as illustrated in Figure 8. Overtopping, gate and breach water volumes were computed using the Weir equation from hydraulic engineering. Water volumes were added to rainfall water volumes and pumping effects were accounted for to produce net water volumes. These water volumes were used in each subbasin to compute water elevation based on respective stage-storage relationships for the subbasins as illustrated in Figure 9.

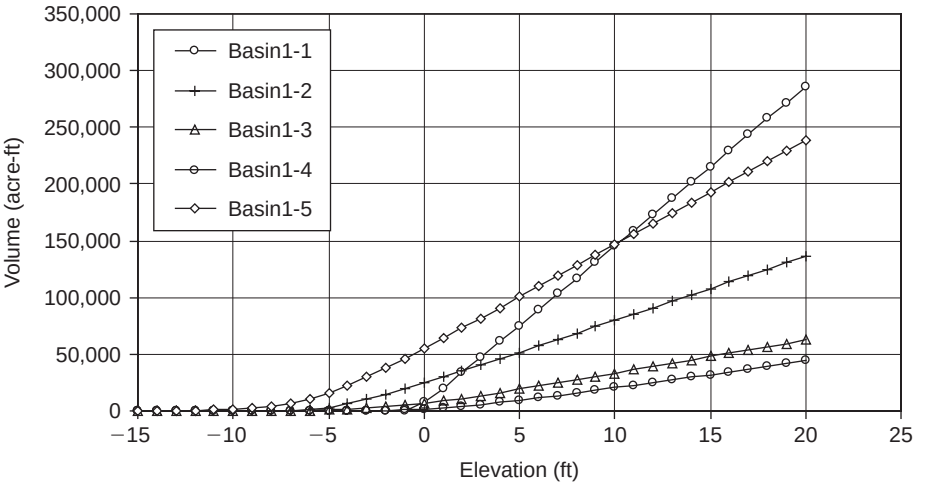


Figure 9 Stage-storage (1 acre-ft = 43,560 ft^3).

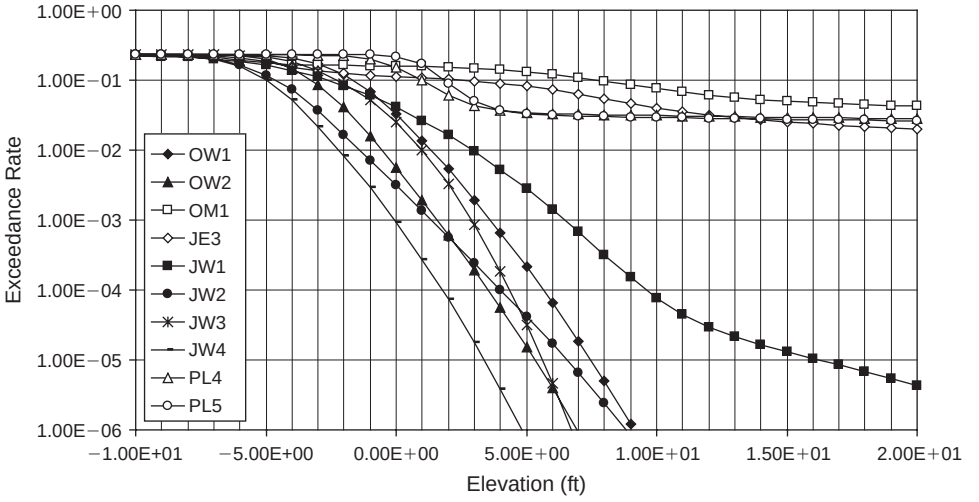


Figure 10 Elevation exceedance curve.

Water interflow logic for subbasins was used to determine final water elevation in each subbasin and compute hazard profiles as elevation-exceedance curves. Figure 10 illustrates such curves. These curves were then combined with loss-elevation curves to produce loss-exceedance curves as illustrated in Figures 11 and 12. Figure 13 shows an example inundation map that can be produced from Figure 9.

Details of this example are provided by Ayyub, et al. (2007a and 2007b), and USACE (2006).

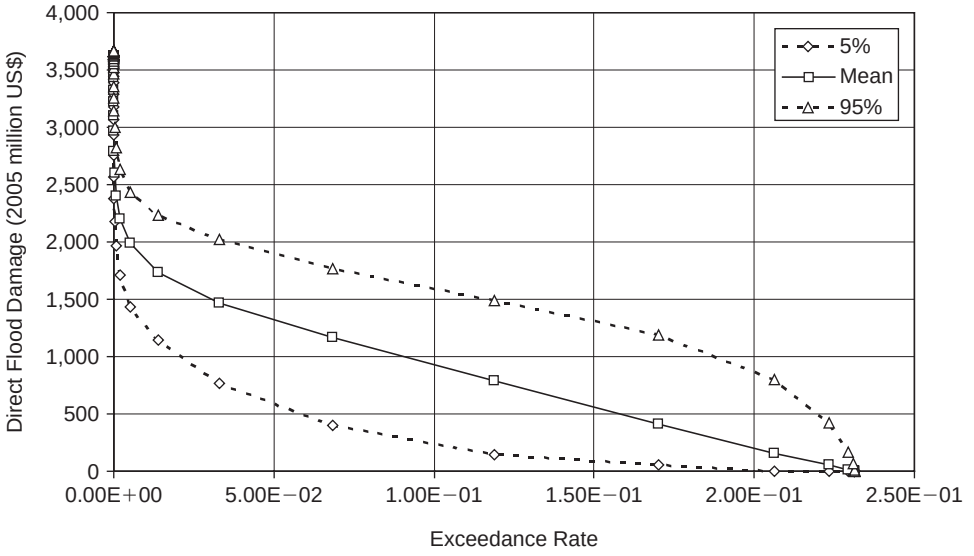


Figure 11 Economic loss exceedance curve.

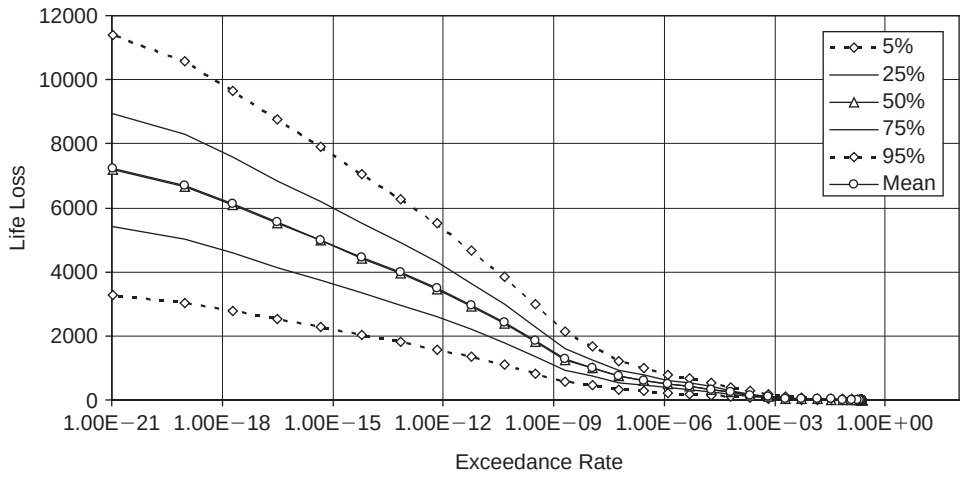


Figure 12 Life loss exceedance curve.

5.2 *All-hazard infrastructure protection*

A five phase process for asset level analysis is used in this example as shown in Figure 14. The five steps are based on the work described in Ayyub, et al. (2007) and McGill, et al. (2007).

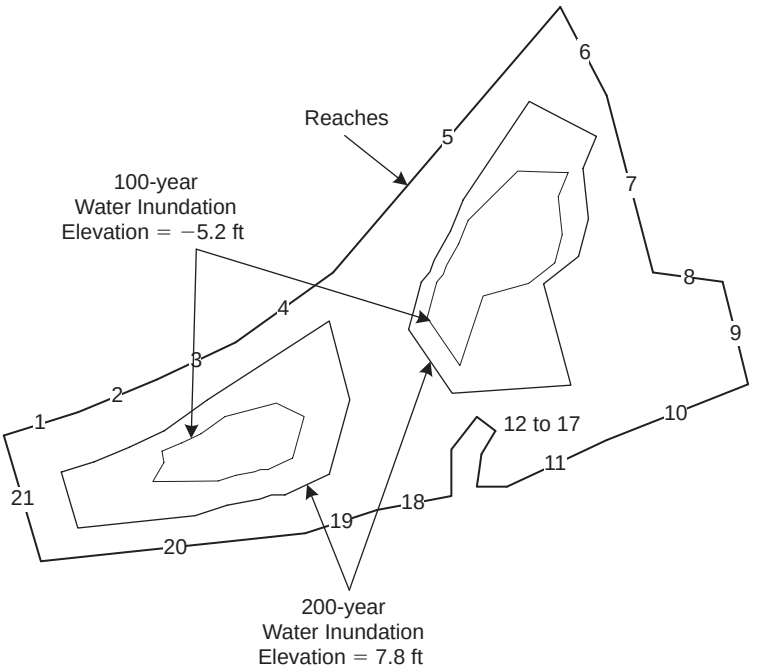


Figure 13 An example inundation map.

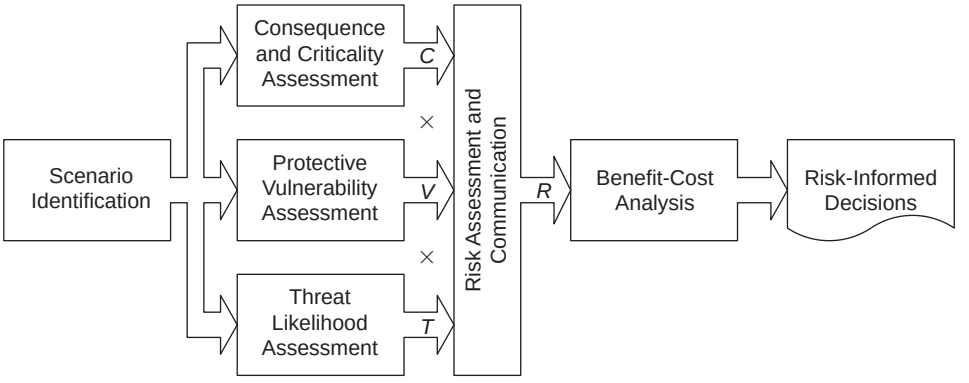


Figure 14 A risk analysis framework.

The scenario identification phase constructs an exhaustive set of hazard and threat scenarios that are relevant to a given asset based on its inherent susceptibilities of its key elements to a wide range of natural and human-caused initiating threat events.

The consequence and criticality assessment phase assesses the loss associated with a given hazard or threat scenario as a function of degree of damage resulting from the damage-inducing mechanisms associated with a hazard or successful attack. The probability of a specified degree of loss L given adversary success at achieving damage

Table 3 Parameter values for risk assessment.

Hazard/Threat	*Maximum credible loss (Millions of dollars)*	*Vulnerability*	*Annual rate of occurrence*
Hurricane	100	0.2 (0.25)	0.2 (0.2)
Tornado	10	0.3 (0.25)	2 (0.2)
Drought	1	0.2 (0.25)	0.1 (0.2)
Winter storm	10	0.01 (0.25)	3 (0.2)
Nuclear attack	500	0.8 (0.25)	1E-06 (0.3)
Explosive attack	3	0.3 (0.25)	0.05 (0.3)
Airplane as projectile	0.5	0.1 (0.25)	0.01 (0.3)
Biological attack	100	0.2 (0.25)	1E-04 (0.3)
Industrial accident	2.5	0.01 (0.25)	0.2 (0.3)

Table 4 Risk assessment results.

Hazard/Threat	*Economic risk (Millions of dollars per year)*	*Sensitivity to changes in vulnerability*
Hurricane	4 (0.32)	0.39
Tornado	6 (0.32)	0.58
Drought	0.02 (0.32)	0.002
Winter storm	0.3 (0.32)	0.03
Nuclear attack	0.0004 (0.39)	3.9E-5
Explosive attack	0.045 (0.39)	0.004
Airplane as projectile	0.0005 (0.39)	4.8E-5
Biological attack	0.002 (0.39)	1.9E-4
Industrial accident	0.005 (0.39)	4.8E-4

SD, $P_{L|SD}$, following the occurrence of a hazard or threat scenario, is computed. With regards to human-caused threats, the protective vulnerability assessment phase assesses the probability of SD for a variety of alternative attack profiles for each threat scenario. An attack profile is the pairing of a specific threat delivery system (such as a vehicle for an explosive attack) with a relevant intrusion path (such as "via main access road"). The effectiveness of measures to detect, delay, respond to, and defeat are considered to arrive at an overall measure of effectiveness, or reliability, of the security or protection apparatus for each attack profile. For human-caused threats, the threat likelihood assessment phase estimates the annual rate of occurrence for each attack profile based on the perceived attractiveness of each the asset, threat scenarios, and attack profiles. The total annual risk expressed as the probability distribution over a continuum of losses L for an asset with respect to a given hazard type.

This example demonstrates a high-level application of the methodology to assess the economic risks to an asset with respect to a full suite of natural and man-made hazards. More details on asset-level examples can be found in the papers by Ayyub, et al. (2007) and McGill, et al. (2007). Note that notional results are used. This example directly assesses the primary parameters in the risk model such as would be the case if limited

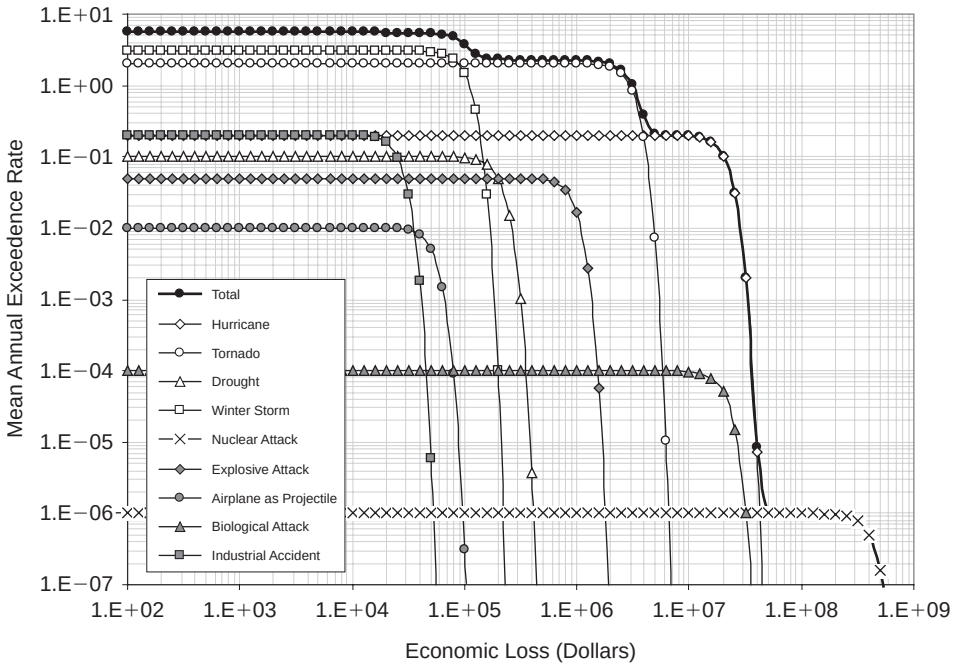


Figure 15 Loss exceedance curves.

resources were available for analysis and data was available primarily in the form of expert judgment.

Table 3 provides an assessment of each parameter and associated coefficient of variation in parentheses. Table 4 gives the contribution to total annual risk from each hazard scenario and the overall sensitivity of risk to a 1% fractional favorable change in the vulnerability. Note that coefficients of variation are given in parentheses adjacent to the mean values. From these results, the total annual risk is 10.4 with a coefficient of variation of 0.22, and improvements in the vulnerability of the region to Tornados should be targeted for cost-effective risk reduction. Moreover, the results from this analysis can be used to construct a family of loss-exceedance curves such as those shown for mean exceedance rate in Figure 15 (Ayyub 2003).

The threat analysis component is not discussed in this example. The reader is directed to the papers by Ayyub, et al. (2007), and McGill, et al. (2007) for additional information.

This example also illustrates the assessment of interdependency losses associated with portfolio level consequence and criticality assessment. Consider a portfolio of three assets – Asset X, Asset Y, and Asset Z – with interdependency matrix $\boldsymbol{K}$ and loss vector $\boldsymbol{c}$ shown in Table 5. Furthermore consider a single hazard type affecting this portfolio, and assume point estimates for the degree of functional degradation and recuperation time for each asset following each hazard event (i.e., hazard afflicting an asset) are given in Table 6. Using the following model and the data from Tables 5 and 6,

Table 5 Portfolio interdependency matrix and daily cost of disruption.

Asset	Percent disruption due to loss of asset (**K**)			Cost per day of disruption (Dollars) (**c**)
	X	Y	Z	
X	NA	0.8	0.3	3,750,000
Y	0.4	NA	0.6	2,500,000
Z	0.9	0.3	NA	1,250,000

Table 6 Resulting interdependency-related loss.

Asset	Service disruption (%/Event) (**u**)	Recuperation time (Days/Event)	Interdependency loss (Dollars/Event)
X	0.6	3	3,825,000
Y	0.2	5	3,375,000
Z	0.5	7	9,187,500

the total interdependency loss for each hazard event was calculated as shown in last column of Table 6:

$$L_P = L_D + L_I \tag{11}$$

where L_D is the direct economic loss (or aggregate loss as appropriate) to the asset assessed from the perspective of the decision maker charged with protecting the portfolio, and L_I gives the loss due to interdependency effects. A simple model for estimating the loss due to interdependency effects can be expressed as:

$$L_I = (\boldsymbol{c}^T \boldsymbol{K} \boldsymbol{u}) L_T \tag{12}$$

where L_T is the time to recuperate lost function following the occurrence a hazard or threat scenario, **c** is a vector that assigns a cost per unit time of disruption for each asset in a given portfolio, $\boldsymbol{K}$ is the portfolio interdependency matrix where elements k_{ij} given the percentage degree of disruption to an asset i due to complete loss of asset j ($k_{ij} = 0$ for $i = j$), and $\boldsymbol{u}$ is a disruption vector whose elements corresponds to the degree of disruption of an attacked asset. Note that the model in Eq. 12 considers only first-order interdependencies, assumes proportional interdependency relationships between assets, neglects substitution, and assumes a proportional relationship between economic loss and degree of disruption per unit time.

Additional details on this example are provided by Ayyub and McGill (2008).

References and bibliography

Ang, A. H-S. and Tang, W. H., 1990. "Probability Concepts in Engineering Planning and Design." Volume II- Decision, Risk, and Reliability. John Wiley & Sons, Inc. New York.

Ayyub, B. M., 2001, Elicitation of Expert Opinions for Uncertainty and Risks, CRC Press, Boca Raton, FL.

Ayyub, B. M., 2003, Risk Analysis in Engineering and Economics, Chapman and Hall/CRC Press, FL.

Ayyub, B. M., and Klir, G. J., 2006, Uncertainty Modeling and Analysis for Engineers and Scientists, Chapman & Hall/CRC, Press Boca Raton, FL.

Ayyub, B. M., and McGill, W. L., 2008. "An All-Hazards Methodology for Ciritical Asset and Portfolio Risk Analysis," ASCE Special Publication *Primer on Quantitative Risk Analysis* (in press).

Ayyub, B. M., Foster, J., McGill, W. L., 2007a. "Risk Analysis of a Protected Hurricane-Prone Region: Model Development," ASCE Natural Hazards Review, in press.

Ayyub, B. M., McGill, W. L., Foster, J., Jones, H. W., 2007b. "Risk Analysis of a Protected Hurricane-Prone Region: Computations and Illustrations," ASCE Natural Hazards Review, in press.

Ayyub, B. M., McGill, W. L., Kaminskiy, M., 2007. "Critical Asset and Portfolio Risk Analysis for Homeland Security: An All-Hazards Framework," Risk Analysis International Journal, Society for Risk Analysis, 27(3), 2007, 789–801.

Ayyub, B.M., and McCuen, R., 2003, Probability, Statistics and Reliability for Engineers and Scientists, Second Edition, Chapman and Hall/CRC Press, Boca Raton, FL.

Kumamoto, H., and Henley, E.J., 1996, Probabilistic Risk Assessment and Management for Engineers and Scientists, Second Edition, IEEE Press, New York.

McGill, W. L., Ayyub, B. M., Kaminskiy, M., 2007. "A Quantitative Asset-Level Risk Assessment and Management Framework for Critical Asset Protection," Risk Analysis International Journal, Society for Risk Analysis, 27(5), 2007, 1265–1281.

Modarres, M., 1993, "What Every Engineer Should Know About Reliability and Analysis," Marcel Dekker, Inc., NY.

Modarres, M., Kaminskiy, M., Krivstov, V., 1999. Reliability Engineering and Risk Analysis: A Practical Guide, Marcel Decker Inc., New York, NY.

USACE, 2006. Interagency Performance Evaluation Task Force Draft Report on "Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System," Draft Volume VIII – Engineering and Operational Risk and Reliability Analysis, USACE, Washington, DC. https://IPET.wes.army.mil

Chapter 12

Resilience and sustainability of infrastructure systems

Masanobu Shinozuka

Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA

ABSTRACT: The primary objective of this paper is to summarize the simulation-based probabilistic methodology for evaluation of performance of spatially distributed systems serving urban population centers under operational and extreme event conditions. The methodology is multidisciplinary involving disciplines of engineering, economics, natural and social sciences. The methodology promotes the system design based on robustness, resilience and sustainability. Critical infrastructures typically include utility and transportation networks which are operationally and functionally interdependent and interactive. The system performance is defined in terms of robustness, resilience and sustainability. This paper focuses on analysis of system robustness and resilience and makes some observations with respect to sustainability. For the purpose of clearly demonstrating the methodology, this paper deals with a model developed for the Los Angeles Department of Water and Power's (LADWP's) power system as part of Western Electricity Coordinating Council (WECC) grid, and the model is used to simulate its robustness and resilience under a set of scenario earthquakes consistent with the regional seismic hazard defined by USGS. The result of the simulation agreed with the robustness and resilience actually demonstrated by the system under the Northridge earthquake. In addition, by employing the model of the entire WECC grid, it is possible to analyze the power flow status within the grid under various component disablements.

1 Introduction

Electric power is essential for virtually every urban and economic function. Failures of electric power networks and grids – whether from natural disaster, technological accident, or man-made disaster such as terrorist attack – can cause severe and widespread societal and economic disruption. In the 1994 Northridge earthquake that struck Los Angeles, some 2.5 million customers lost electric power. For the first time in its history, the entire city of Los Angeles was blacked out. Furthermore, power outages were experienced in many areas of the western U.S. outside the seismically affected region and as far away as Canada (Hall, 1995). On August 14, 2003, blackout of unprecedented proportions rippled out from Akron, Ohio, across the northeastern U.S. and parts of Canada, affecting an area with a population of some 50 million (U.S.-Canada Power System Outage Task Force, 2003). In September of 2003, a power outage that began in Switzerland cascaded over a large region of Italy. Examples such as these indicate the importance of being able to anticipate potential power system failures and identify effective mitigation strategies.

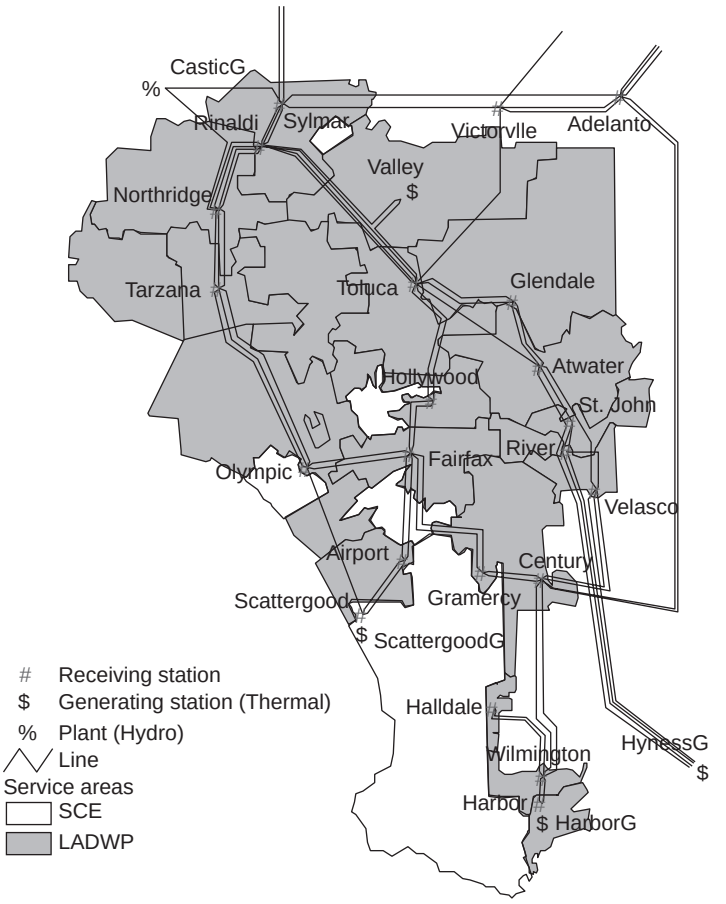


Figure 1 Transmission Network and Service Areas of LADWP's Power System.

Modeling the impacts of electric power disruption is, however, a highly complex problem. Many of the inherent challenges relate to the need to integrate across disciplines – not only civil, mechanical, and electrical engineering, but also economics and other social science disciplines. For example, one must assess how damage to individual pieces of electric power equipment affects power flow across the network. One must model how a damaged network would be repaired and how electric power would be restored over space and time. Additionally, one must capture how the loss of electric power would affect businesses and other units of society.

While numerous studies have addressed portions of the problem, very few have attempted to model regional impacts. For example, Nojima and Sugito, 2003 modeled power restoration times based on empirical data from Japanese earthquakes using geographical information systems (GIS); however, they did not evaluate power outage and its impacts on actual urban areas. The current study builds on a long-standing research program led by the author and carried out by a group of researchers. Their results

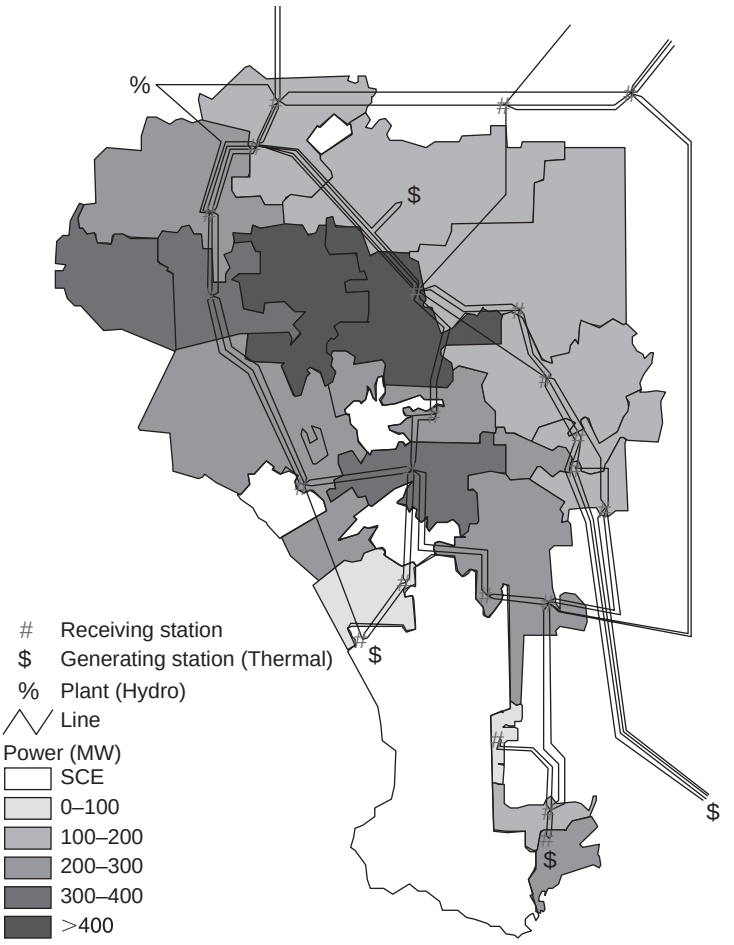


Figure 2 Electric Power Output for LADWP's Service Areas under Intact Condition.

(e.g., Shinozuka et al., 2002 and Shinozuka and Dong, 2002, Tanaka et al., 1997) laid the foundation for the development of a method by which utility transmission network performance can be rationally analyzed, taking into account power flow in a seismically damaged network. Related research developed methods for assessing how power system failures would impact on regional economies (Rose et al., 1997).

The objective of this paper is to demonstrate the method of evaluating the seismic robustness and resilience of electric power system, to recommend appropriate seismic rehabilitation measures and to assess socio-economic impacts. The Los Angeles Department of Water & Power's (LADWP's) power system was taken as a test-bed in this demonstration. Figs. 1 and 2 show LADWP's electric power service areas and the power supply under normal operating conditions. The areas not colored are serviced by Southern California Edison (SCE). Fig. 3 is the distribution of the customers who are power-supplied by LADWP. Fig. 4 is the distribution of day-time population who

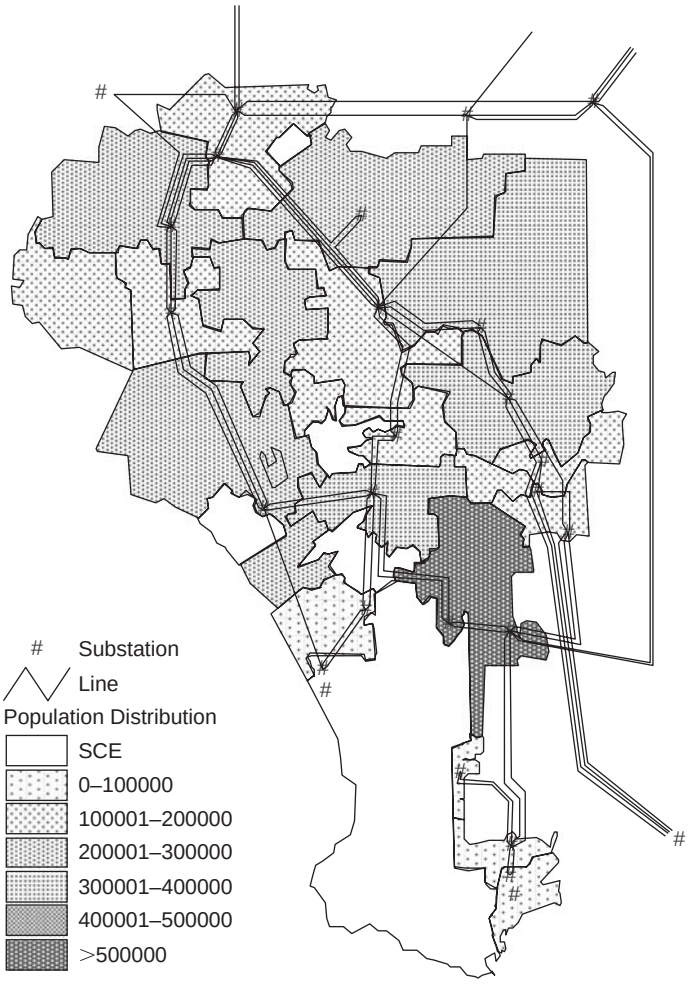


Figure 3 Distribution of Population Supported by LADWP.

are power-supplied by LADWP. Fig. 5 is the distribution of the customers who are power-supplied by LADWP. Fig. 6 is the distribution of hospitals which are power-supplied by LADWP. To develop the analysis methodology and gain insight into the performance of the power system, fragility curves of electrical power equipment, such as transformers in the transmission network, were developed on the basis of damage information from the 1994 Northridge earthquake. This paper also uses results from an inventory survey and equipment rehabilitation study being performed concurrently by members of the research team, supported by the Multidisciplinary Center for Earthquake Engineering Research (MCEER). With the aid of systems analysis procedures, a performance analysis of LADWP's power system was conducted for actual and simulated earthquakes, using a network inventory database, available fragility information,

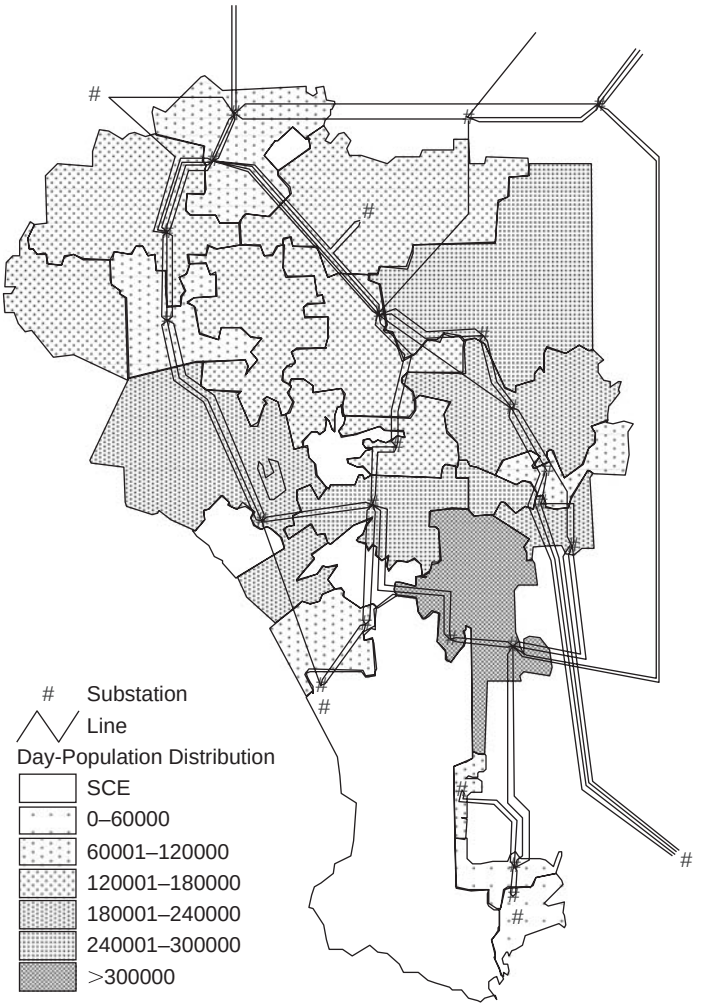


Figure 4 Distribution of Day-Time Population Supported by LADWP.

and Monte Carlo simulation techniques. This is a unique research work in which the Western Electricity Coordinating Council's (WECC) database are used for the systems analysis, in conjunction with the computer code IPFLOW (version 5.2), licensed by the Electric Power Research Institute (EPRI).

In addition to transformers, the seismic vulnerability of other equipment and components such as circuit breakers disconnect switches and buses, and their impact on system performance is integrated into the analysis by representing their vulnerability in terms of respective fragility curves.

To gain a more complete understanding of the performance of LADWP's power system under the possible seismic action in the study area, 47 scenario earthquake

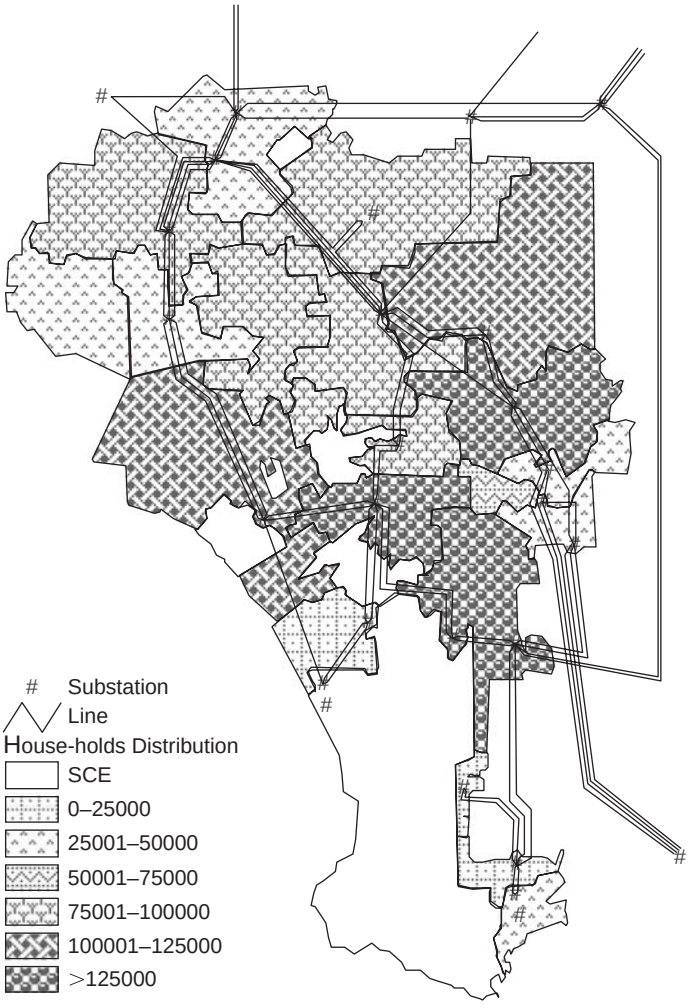


Figure 5 Distribution of Customers Supported by LADWP.

events were selected and simulated. By including each scenario's associated annual "equivalent probabilities" of occurrence, they represent the full range of the regional seismic hazard curve (WECC website). Based on the power flow analysis results from these 47 events, the risk curves for system performance degradation, for example, reduction in power supply, and in GRP (Gross Regional Product) immediately after an earthquake in LADWP's service areas were developed.

A repair and restoration model was developed to predict and evaluate the restoration process of the power systems. The system restoration process was simulated and restoration curves were developed from the results of the power flow analysis.

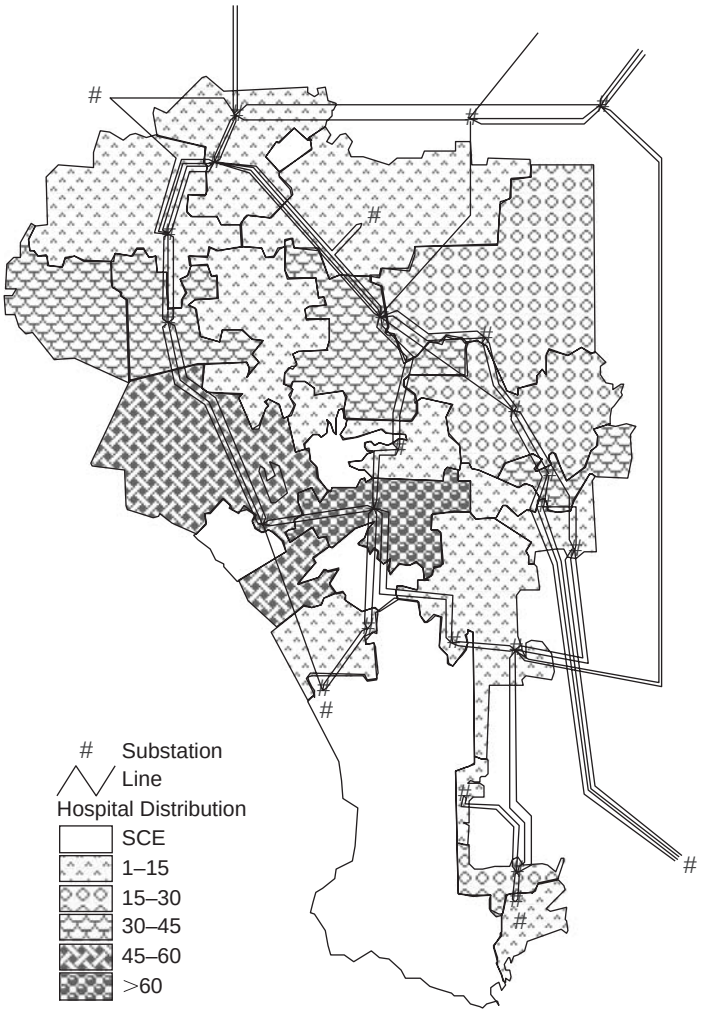


Figure 6 Distribution of Hospitals Supported by LADWP.

2 Technical summary

2.1 *Seismic performance of Ladwp's power system*

2.1.1 *Transmission systems*

A utility power system consists of generating stations, transmission systems and distribution network. The current paper focuses on transmission systems including receiving stations. Furthermore, an assumption is made so that the transmission lines will not fail under seismic condition. This assumption is generally acceptable for LADWP's system and allows one to concentrate on receiving stations. There are many electric components in receiving stations, such as transformers, circuit breakers, disconnect

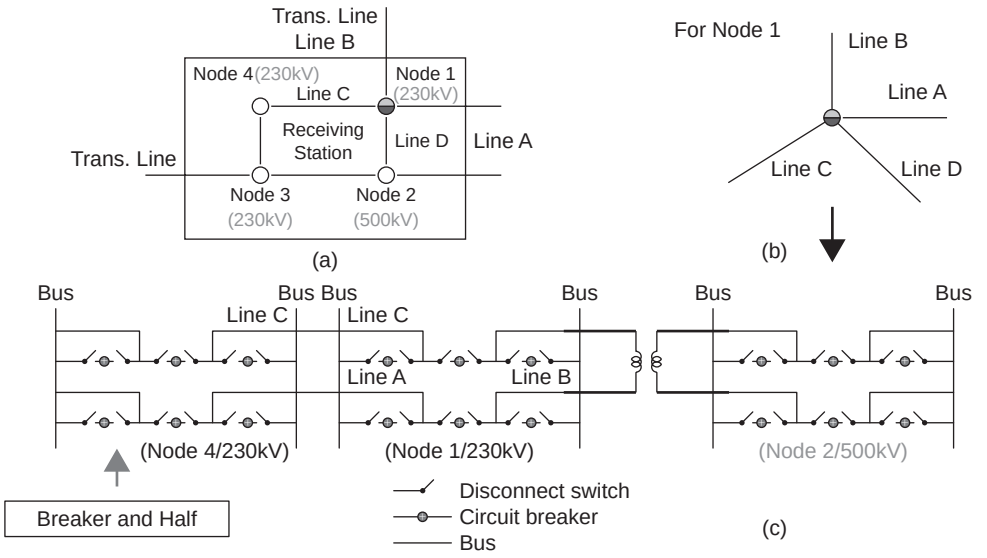


Figure 7 Models for receiving station and node.

switches, lightening arresters, current transformers, coupling voltage transformers, potential transformer, wave trap and circuit switches. These components are connected to transmission lines through buses. Transmission lines then connect links between generating stations and distribution systems and lead to other power systems. In general, if the voltage between two buses is different, then there must be at least one transformer between them. Fig. 7 is the models for substation and nodes. Fig. 7(a) is a model of substation with four nodes and Fig. 7(b) depicts a node connected to transmission lines connected to it. A node, which consists of buses, circuit breakers and disconnected switches, can take many configurations to increase the level of redundancy so as to minimize the chance of the transmission lines disconnected and hence to maximize the reliability of the node as well as the reliability of electric power system as a whole. Actually equipment configuration in a node is complex. A general popular node configuration is known as a "breaker and a half" model as shown in Fig. 7(c).

2.1.2 *Seismic performance of power system*

LADWP's network is part of the very large WECC power transmission network, covering 14 western states in the USA, two Canadian provinces in Canada and northern Baja California in Mexico. The present analysis considers 52 receiving stations within the WECC network that are subjected to significant ground motion intensity under the 47 scenario earthquakes. Using ArcGIS platform, the map of 52 receiving stations in Fig. 8 is overlaid on the map of peak ground acceleration (PGA) from the 1994 Northridge earthquake as shown in Fig. 9(a) to identify the PGA value at the location of each receiving station. The fragility curves provided in Fig. 10 were then used to simulate the damage state for transformers at each of the 52 receiving stations (some

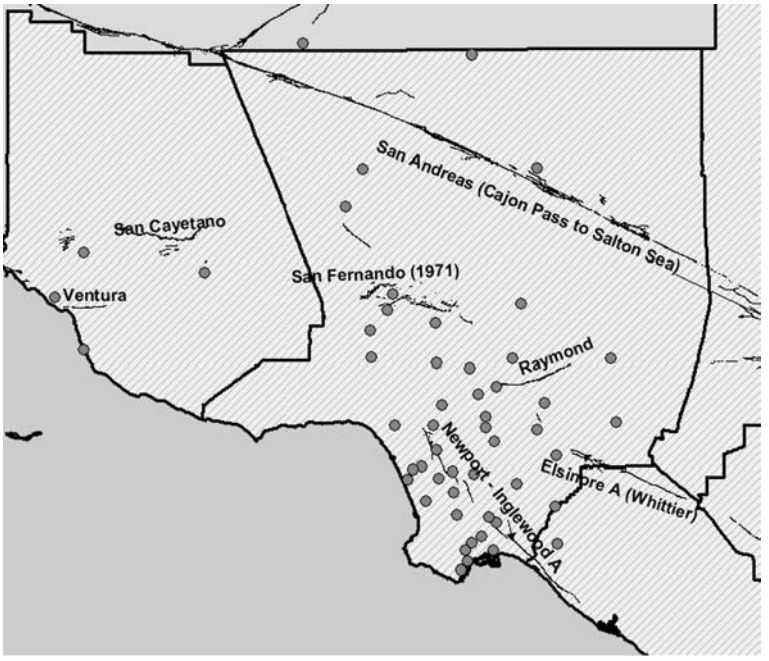


Figure 8 Locations of Earthquake Faults and 52 Receiving Stations.

in LADWP and others in SCE power systems). Note that 3 fragility curves (labeled Case 1, 2 and 3) are given in Fig. 10 where Case 1 curve is obtained empirically from the Northridge earthquake damage data, Case 2 curve represents improvement on Case 1 curve by 50% (in terms of median value) and Case 3 curve by 100%. These improvements are deemed possible on the basis of analytical and experimental study by Saadeghvaziri and Feng, 2001. For each system analysis, connectivity and power flow were examined with the aid of IPFLOW, where LADWP's power system was treated as part of the overall WECC system.

The methodology used to evaluate the seismic performance of the electric power network is described in the following steps as also described in the flow chart in Fig. 11.

(1) Use 47 scenario earthquakes (WECC website) (13 Maximum Credible Earthquakes and 34 User Defined Earthquakes)
(2) For each scenario earthquake, simulate equipment damage using fragility curves representing conditions with and without rehabilitation
(3) Simulate damage to the transmission network
(4) Calculate power flow using IPFLOW under the network failure criteria

(i) Imbalance of power:

$$1.05 > \frac{total\ supply}{total\ demand} \text{ or } \frac{total\ supply}{total\ demand} > 1.1 \qquad (1)$$

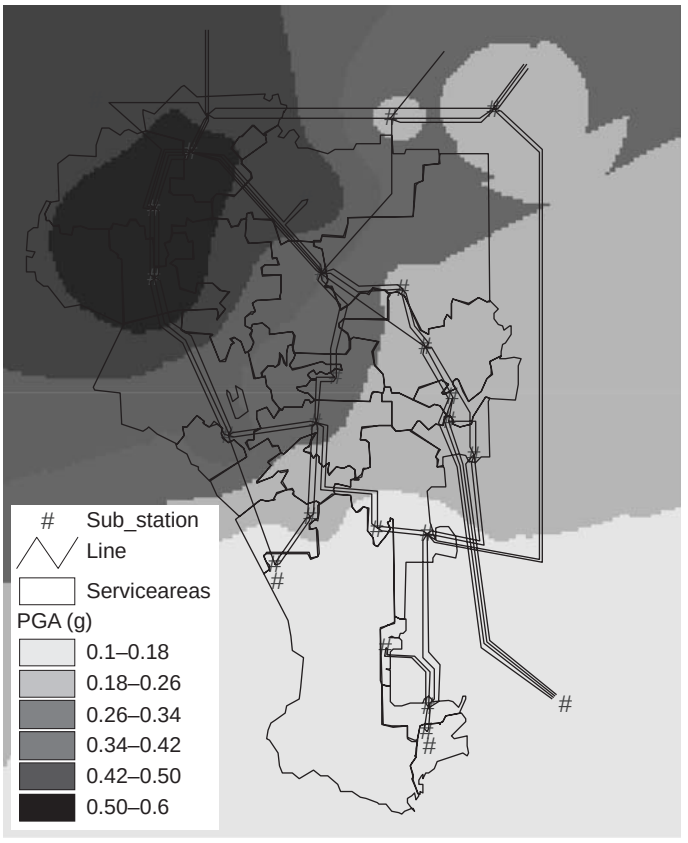


Figure 9(a) Spatial PGA Distribution in the 1994 Northridge Earthquake.

(ii) Abnormal voltage

$$\left| \frac{V_{\text{int } act} - V_{damage}}{V_{\text{int } act}} \right| > 0.1 \tag{2}$$

(iii) Frequency change (IPFLOW does not check this criterion)
(iv) Loss of connectivity

(5) Compute the seismic performance of the power network in terms of percentage of power supply and number of customers in the entire area of service as well as each service area under each scenario earthquake
(6) Evaluate reduction in the seismic performance
(7) Develop seismic risk curve (which plots the annual probability that system performance will be reduced more than a specified level due to earthquake as a function of that level)
(8) Examine system performance relative to performance criteria, with and without rehabilitation
(9) Determine the effectiveness of rehabilitation

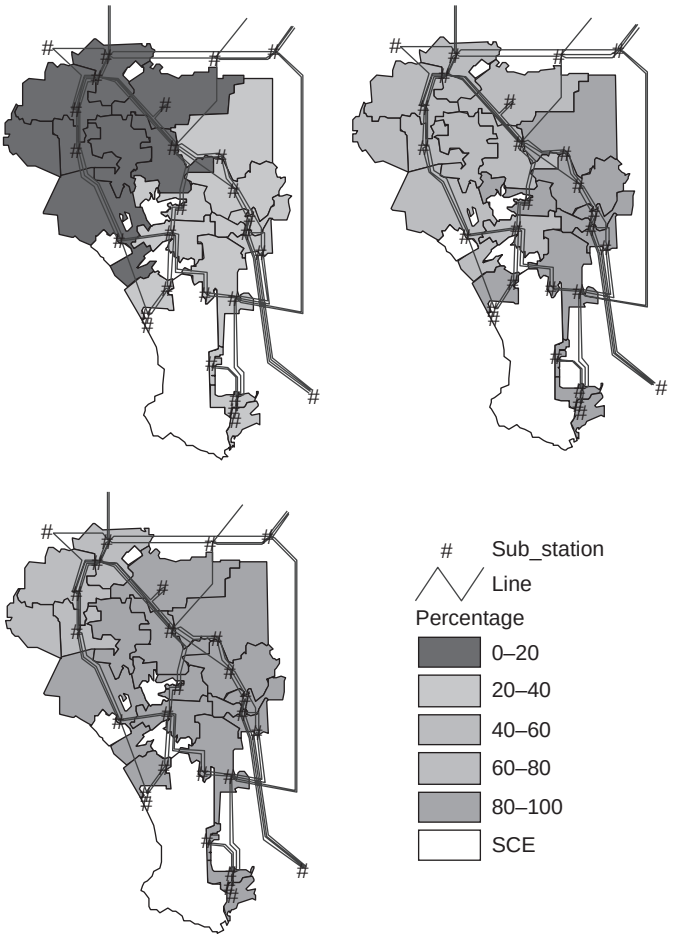


Figure 9(b) Relative Average Power Output with only Transformers Assumed to be Vulnerable (Sample size = 20).

Utilizing Monte Carlo simulation techniques involving the fragility curves in Fig. 10, the power flow analysis is performed on LADWP's network 20 times under each scenario earthquake. Each simulation result represents a unique state of network damage. Fig. 9(b) shows the ratio in percentage of the average power supply of the damaged network to that associated with the intact network for each service area, when only transformers are considered to be vulnerable. The average is taken over all simulations. The extent to which the rehabilitation of transformers contributes to improvement of system performance is evident if we compare the power supply ratio under Case 1 (not enhanced), Case 2 (50% enhanced) and Case 3 (100% enhanced) as shown in Fig. 9(b).

In addition to transformers, there are other important types of electrical equipment in the power network that are vulnerable to earthquake ground motion. They include circuit breakers and disconnect switches and buses.

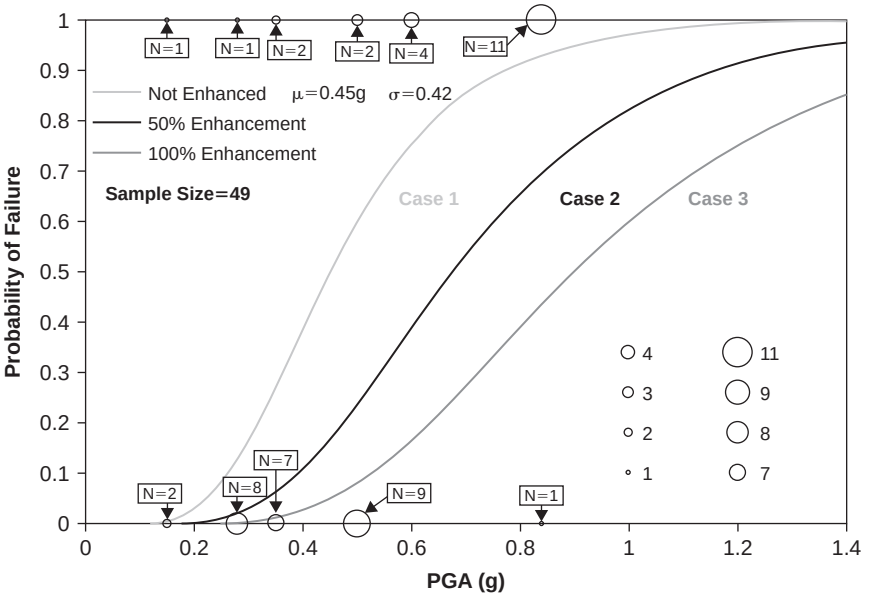


Figure 10 Fragility Curves for Transformers with and Without Enhancement.

2.2 *Risk evaluation of power systems*

2.2.1 *Scenario earthquakes*

For electric power and other urban infrastructure systems, evaluating potential impacts of damage is complicated by the fact that the networks are spatially distributed over a wide area. Risk analysis must account for how the system performs given that the hazard (e.g., earthquake ground motion) is not only spatially variant across a wide area but also, for any given disaster, spatially correlated. Hence, traditional probabilistic methods that can readily be applied for site-specific facilities such as individual buildings cannot be used for these spatially distributed networks.

The current methodology therefore analyzes system functionality and impacts in the context of scenarios of individual earthquake events, then combines the scenario results probabilistically to gain a complete understanding of the seismic performance of LADWP's power system. This is done in the form of risk curves that graphically summarize system risk in terms of the likelihood of experiencing different levels of performance degradation in disasters. Risk curves can be developed for performance parameters associated with different dimensions of robustness, including the technical (e.g., power supply in each service area), societal (e.g., rate of customers without power supply), organizational (e.g., repair and restoration efficiency), and economic (e.g., regional output or employment loss). The following GIS database for population (Fig. 3), day population (Fig. 4), customers (Fig. 5) and hospitals (Fig. 6) are used to develop these risk curves.

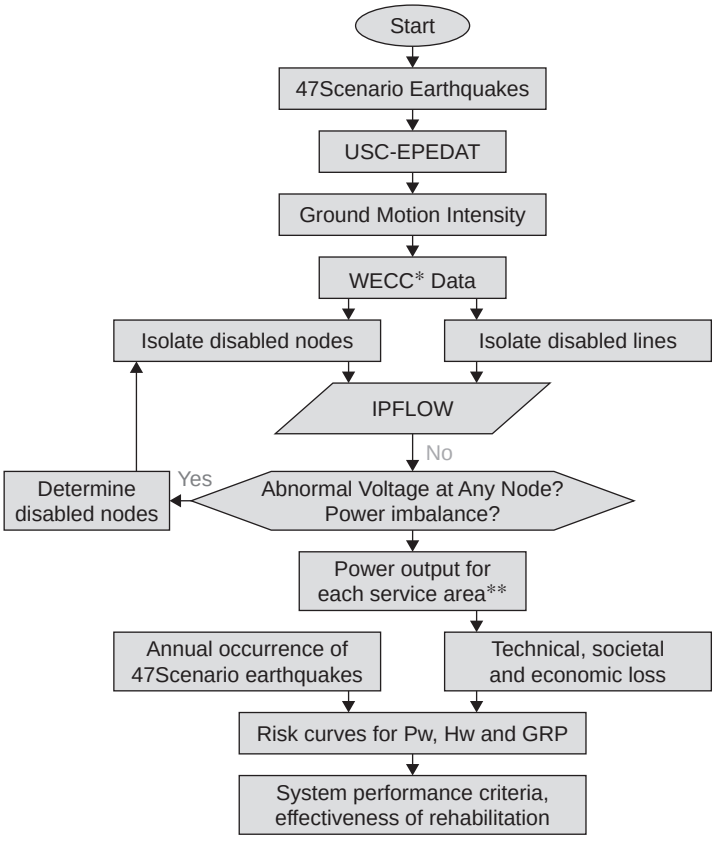


Figure 11 Power Performance Analysis Flow Chart.

2.2.2 *Economic impact*

The preceding analysis of systems performance can be readily extended from impacts on power supply capability to impacts on the regional economy. Here, direct economic losses are evaluated using a methodology that relates the spatial pattern of electric power outage to the regional distribution of economic activity (Shinozuka and Chang, 2004).

Direct economic loss, L (dollars), is evaluated for each earthquake simulation and each mitigation condition as follows:

$$L = \sum_{s} \sum_{j} l_j \cdot d_s \cdot e_{s,j} \tag{3}$$

where l_j is a loss factor for industry j ($0 \leq l \leq 1$), d_s is a disruption indicator for service area s ($d = 1$ in case of power outage, $d = 0$ in case of no outage), and $e_{s,j}$ is daily

industry j economic activity in area s (dollars). The disruption indicators d_s for each electric power service area derive directly from the power supply simulation results described previously.

The loss factors l_j reflect the dependency of each industry on electric power. They were developed empirically on the basis of survey data collected following the 1994 Northridge earthquake that struck the Los Angeles region. Specifically, a large survey of over 1100 businesses was conducted by K. Tierney and colleagues at the Disaster Research Center of the University of Delaware (see Webb et al., 2000). Data from this survey that were used in the current study included information on whether a business lost electric power, for how long, the level of disruptiveness associated with this outage, and whether or not the business closed temporarily in the disaster. Data on other sources of disruption (e.g., building damage, loss of water, etc.) were also used to estimate the net effect of electric power outage. For details on the methodology, see Chang et al., 2002 and Chang and Seligson, 2003. The loss factors range from a low of 0.39 for mining and construction to a high of 0.60 for manufacturing. These factors pertain to a one-day power outage.

Estimates of industry economic activity by service area, $e_{s,j}$, were based on industry employment data. Employment by industry and zip code were obtained from the Southern California Association of Governments (SCAG) and aggregated, using GIS overlays, to the LADWP service areas. Employment was converted into output using estimates of output per employee in each industry. These productivity estimates were based on California gross state product (GSP) and employment data available from the Bureau of Economic Analysis (BEA).

Loss results are expressed as the percent of gross regional product (GRP) in the LADWP service area that would be lost given electric power outage in each earthquake simulation. Because the duration of power outage is not modeled here, results are assessed in terms of daily GRP loss. This can be interpreted as the loss that would obtain if the outage pattern lasted for one day.

2.2.3 *Risk curves for LADWP's power system*

The locations of LADWP and SCE receiving stations relative to faults in and around the Los Angeles area indicate that many are near active faults (Fig. 8) and have a high likelihood of suffering from damage due to an earthquake. To evaluate the risk and costs associated with potential future earthquakes, the performance of the power system in 47 (deterministic) earthquake scenarios was evaluated, as noted earlier. Based on these scenarios, together with their hazard-consistent probabilities, the seismic risk to LADWP's power system performance due to earthquake-induced performance degradation is evaluated in terms of "risk curves." Figs. 12 and 13 shows the risk curves in terms of the severity of power loss for Case 1 (no fragility enhancement) and for Cases 2 and 3 (enhancement index 50% and 100% respectively) for reduction in power supply, and reduction in GRP. One can develop similar risk curves for each service area. The significant improvement due to seismic retrofit is clearly seen in these risk curves. The following is mathematical expression to used for development of risk curves as a function of percentage loss of power supply. Similar expression is used for the development of risk curves for GRP.

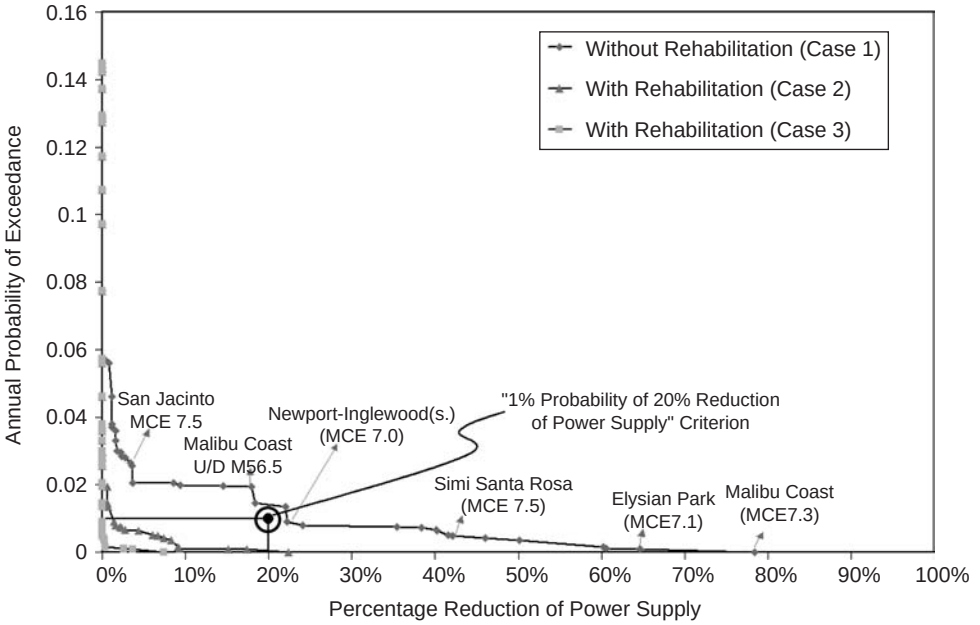


Figure 12 Risk Curves for Power Supply Reduction.

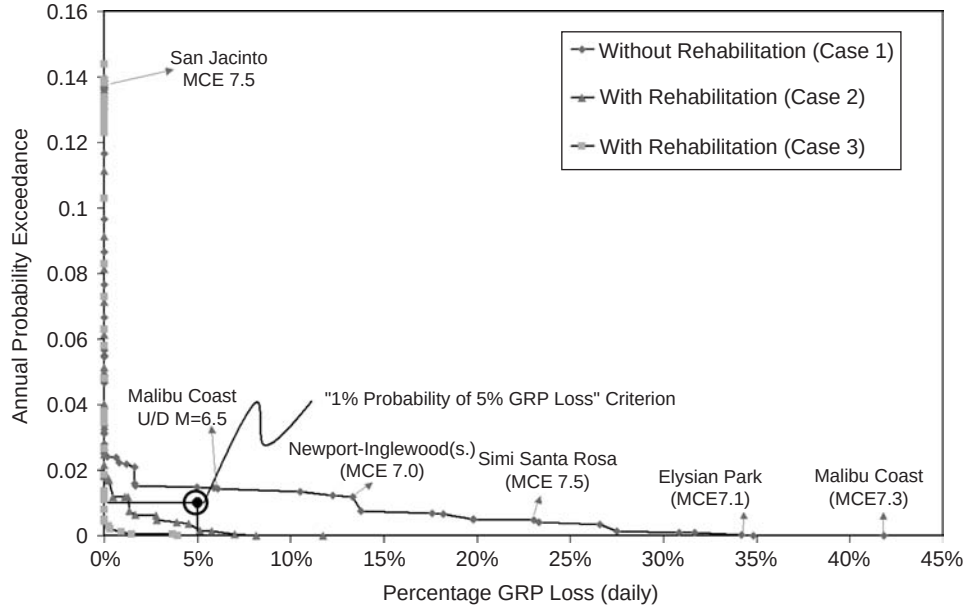


Figure 13 Risk Curves for Economic Loss.

Table 1 System Performance Criterion I for Pre-event Assessment and Rehabilitation.

	Robustness	*Reliability*
Power	A majority (at least 80%) of customers will have continued power supply after earthquake	With a high level of reliability (at least 99% per year)
Water	A majority (at least 80%) of customers will have continued water supply after earthquake	With a high level of reliability (at least 99% per year)
Hospital	A majority (at least 95%) of injured or otherwise traumatized individuals will be accommodated in acute care hospitals for medical care	With a high level of reliability (at least 99% per year)

Percentage of power supply P_w

$$P_w = \frac{\sum_{m=1}^{M} \frac{1}{N} \sum_{n=1}^{N} Pd(m,n)}{\sum_{m=1}^{M} P(m)} \tag{4}$$

Percentage of reduction in power supply P_{wo}

$$P_{wo} = 1 - P_w \tag{5}$$

where:
m = service area number (1,2,. . ., M); M = 21 in this example
n = simulation number (1,2,. . ., N); N = 20 in this example
Pd(m,n) = power output in service area m under n-th simulation
P(m) = power output in service area m under normal condition simulation
The risk curve approach is useful for economic impact analysis, as well as cost-benefit analysis to determine the effectiveness of enhancement technologies (see the curves with solid triangles and squares) (Dong, 2002). Of equal importance is the use of the risk curve in relation to the development of performance criteria and their verification.

2.3 *System performance criteria*

The performance criteria for power systems listed in Tables 1 and 2 demonstrate a possible format in which the criteria can be given. Table 1 lists criteria to be satisfied in pre-event assessment (e.g., through seismic retrofit), and Table 2, those in post-event emergency response (e.g., through disaster response planning). These tables also include performance criteria for water and acute care hospital systems. This general format for performance criteria for structures and lifelines has been provided by Shinozuka and Dong, 2002 and Bruneau et al., 2003. In combination, they conceptually establish the degree of community resilience in terms of robustness, rapidity and reliability. Specific values (in percentages for robustness, rapidity in restoration, and reliability) are examples so that the concept can be better understood. The performance criterion for power systems shown in Table 1 is represented by a point double-circled in Fig. 12

Table 2 System Performance Criterion II for Post-Event Response and Recovery.

	Rapidity in Restoration	*Reliability*
Power	A majority (at least 95%) of customers will have power supply as rapidly as possible within a short period of time (3 days)	With a high level of reliability (at least 90% of earthquake events)
Water	A majority (at least 95%) of customers will have water supply as rapidly as possible within a short period of time (3 days)	With a high level of reliability (at least 90% of earthquake events)
Hospital	All the injured and traumatized individuals will be accommodated in acute care hospitals as rapidly as possible within a short period of time (1 day)	With a high level of reliability (at least 90% of earthquake events)

where the robustness corresponds to the annual probability of 0.01 that 20% more reduction in power supply will result immediately after any earthquake.

Data collection and modeling for rapidity in restoration are much more difficult to pursue (Shinozuka et al., 2002). Further research is needed to develop analytical models based on past experience so that performance criteria, such as those shown in Table 2, become meaningful in practice. However, some simulation study was performed and compared with the Northridge repair/restoration data. The result of this study provides a potentially successful method of pursuit in this area as demonstrated below in section of "System Restoration."

2.4 *Resilience framework and system restoration*

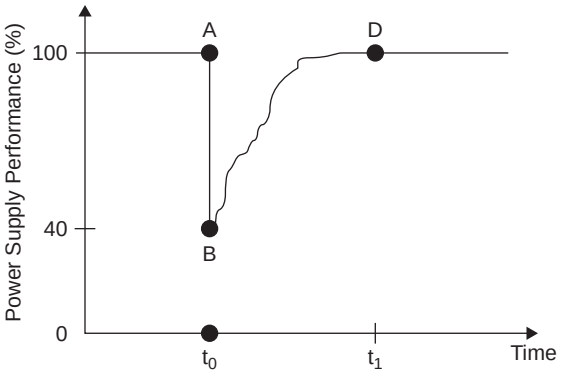
Resilience is an important concept for the disaster management of infrastructure systems. Two key dimensions of resilience can be referred to as robustness and rapidity in restoration. These can be expressed utilizing a restoration curve typically having characteristics as shown in Fig. 14.

The curve plots system performance as a function of time. The reduction in performance from 100% at point A (time t_0) to 40% (in this example) at point B results from the damaging seismic impact to the system. The restoration curve starting from the initial distress point B, to the complete recovery point D (back to 100% at time t_1), demonstrates the process of restoration. Hence, the performance percentage corresponding to B (or B-C, with C associated with zero performance) represents robustness (Equation 6), and the elapsed time for the total restoration (t1-t0) can be used to quantify rapidity (Equation 7), although Equation 7 may admittedly be too simplistic.

$$\text{Robustness} = B - C \text{ (in percentage)} \tag{6}$$

$$\text{Rapidity} = \frac{A - B}{t_1 - t_0} \text{ (recovery in percentage/time)} \tag{7}$$

It has been demonstrated that the restoration for power systems tends to be rapid compared with that for water, gas and transportation systems. Fig. 15 shows the



t_0 = Time at which an earthquake occurs

t_1 = Time at which power supply is restored 10%

Figure 14 Power Supply as a Function of Time.

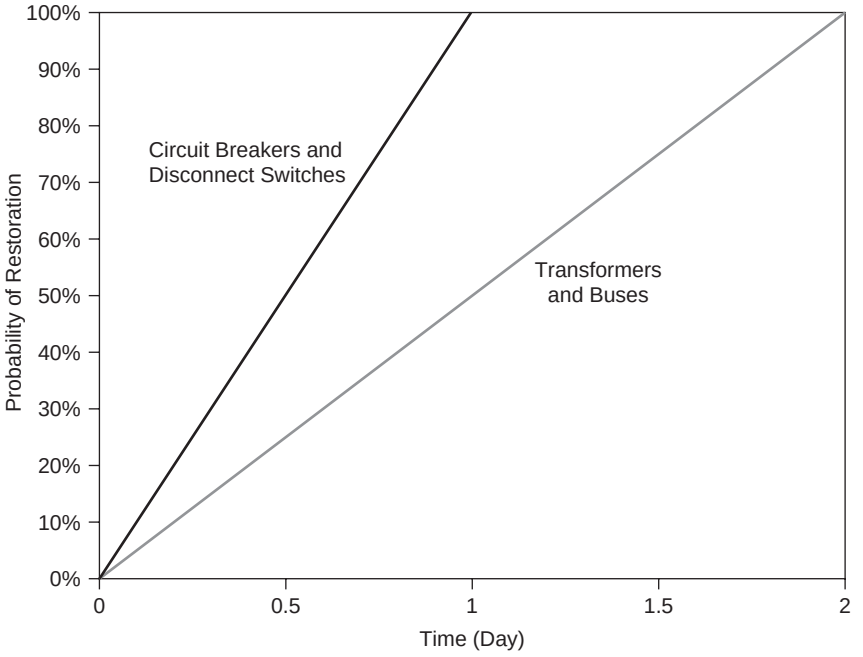


Figure 15 Restoration Curve for Transformers, Circuit Breakers, and Disconnect Switches.

assumed repair or replacement curves for the LADWP system after the Northridge earthquake. The curve plots the probability of damaged equipment being restored (repaired or replaced) as a function of time (in days). It is postulated that circuit breakers and disconnect switches are more rapidly restored with uniformly distributed

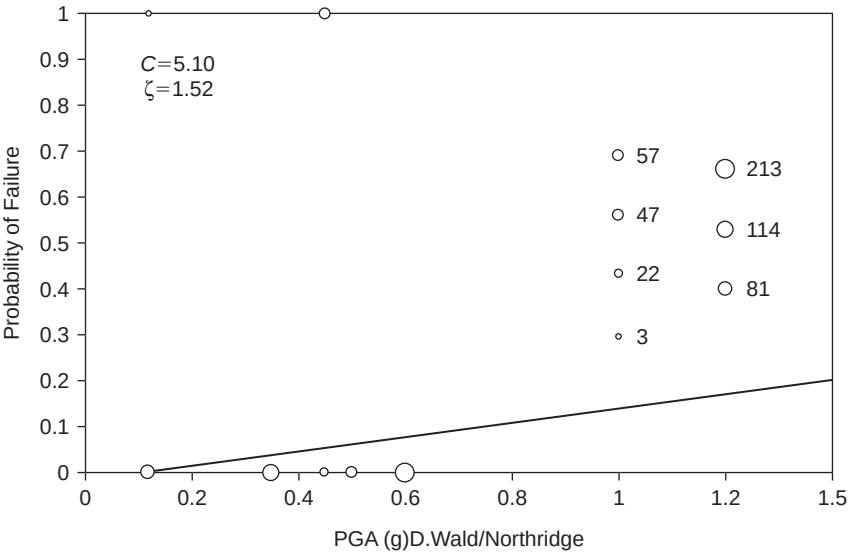


Figure 16 Fragility Curve for Circuit Breakers.

probability density over the first one day period, and transformers and buses over the first two days. This not only reflects the relative ease with which each component is repaired/replaced but also the cost of its replacement. The resulting curve indicates for example that a damaged transformer can be replaced or repaired within a half day with a probability of 25%. This is merely an assumption on which we initiate and gain numerical insight for the restoration simulation. In reality, a transformer probably cannot be replaced or repaired with such rapidity unless the degree of damage is moderate. We follow the fragility simulation, Fig. 10 (Case 1) for transformers Fig. 16 for circuit breakers and Fig. 17 for disconnect switches and buses. The fragility curves for circuit breakers and disconnect switches are also developed from the Northridge damage data. Then, power flow analysis as outlined in earlier sections is performed and another layer of Monte Carlo simulation is added where damaged components are restored in accordance with the restoration probability function assumed in Fig. 15. As for the fragility information, The resulting simulation of restoration (in % of customers with power) is represented at ½, 1, 1½, 2 days after the earthquake in Fig. 18, which somewhat underestimates the speed of restoration actually observed in the aftermath of the Northridge Earthquake. Fig. 18 also includes simulated restoration curves for two other less damaging scenario earthquakes. Note their shapes are, in essence, the same as the curve BD in Fig. 14. We note that power system restoration procedures may repair or replace damaged components in an order reflecting the priority established by the utility for the purpose of accelerating the entire network restoration. If such a procedure were taken into consideration, the simulation performed here might have more closely agreed with the empirical curve. Fig. 19 combines the risk curves and restoration curves under various scenario earthquakes (6 of them in this case) and

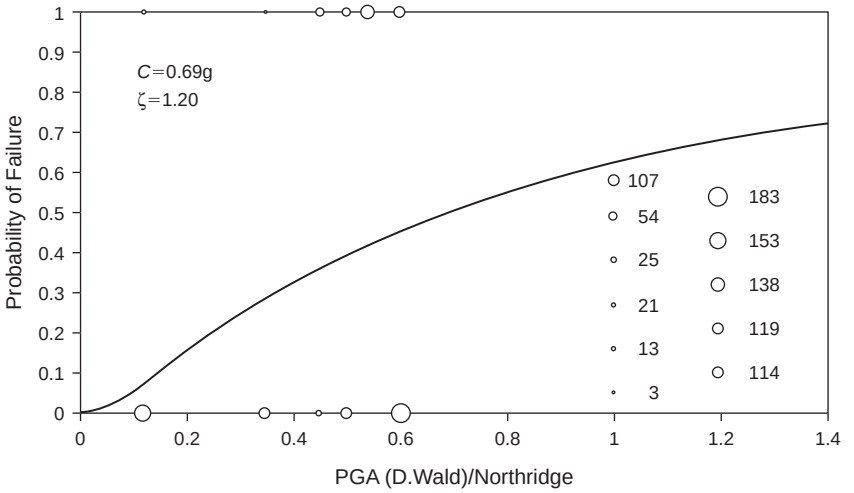


Figure 17 Fragility Curve for Disconnect Switches and Buses.

depicts the robustness and resilience of the LADWP's transmission system under the seismic hazard as determined by USGS. The simulated states of restoration as time proceeds can be depicted in GIS format. Fig. 20(a) shows a snapshot at 6 hours after the earthquake of this spatio-temporal progression of restoration process as reported by LADWP, whereas Fig. 20(b) shows the simulated version of the state of restoration at the same time. The agreement is extremely good and indicates the validity of all the models and parameters used for simulation.

2.5 *Sustainability*

There is no simple definition for sustainability currently available for civil infrastructure systems, at the risk of over simplification, we may define sustainability in such a way that a system is sustainable under natural, accidental, and manmade hazards, if it is designed sufficiently resilient relative to the return period of the extreme events arising from these hazards. Sustainability depends on complex interactions of technical, economical, societal and geographical issues. Typically Fig. 21 shows yearly throughput productivity at Kobe port. The port regained a respectable productivity after Kobe earthquake (1995). However, its world ranking continues to slide from #5 before the earthquake to #35 in 2005 indicating a complex nature of shipping business being affected by the interruption of port operation after the earthquake.

3 Accidental disablements

Fig. 22 shows the impact of a virtual accident that occurred in the State of Oregon in the form of disablement of a segment of 50 kV Western grid transmission line. The impact cascades through the western grid. Even Los Angeles system is affected but quite differently depending on the location of disablement. In one case, the disablement

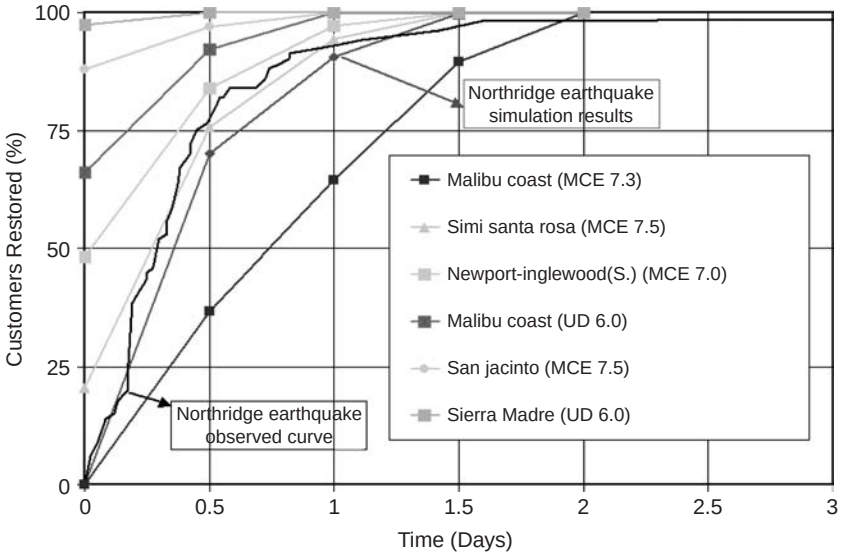


Figure 18 LADWP Power System Customers Restoration.

occurred on a more redundant segment (Fig. 22a; partial blackout) than the other location on a non-redundant segment (Fig. 22b; total blackout). This disablement scenario is the same as the one that caused a great north-east blackout in the summer of 2003. It is important to note that this disaster scenario can result from terrorist disabling a segment.

4 Conclusions

Significant original achievements are summarized in this paper that will remain lasting conceptual and analytical tools for performance prediction of electric power systems in particular and lifelines in general under uncertain hazard conditions. They are listed below.

1. Developed and used a power flow analysis package integrating WECC database and IPFLOW computer code for performance evaluation of LADWP system and Western Grid under natural and other hazards.
2. Introduced a method of developing a set of scenario earthquakes consistent with the USGS seismic risk.
3. Defined system performance as robustness (residual supply capability) based on power flow analysis after a seismic event.
4. Recognized risk curves as representing the robustness of spatially distributed infrastructure systems subjected to seismic hazards.
5. Introduced the predictive model for seismic performance of infrastructure systems by integrating risk curves and restoration curves
6. Developed a seismic damageability of receiving stations using “breaker and a half” bus configuration model.

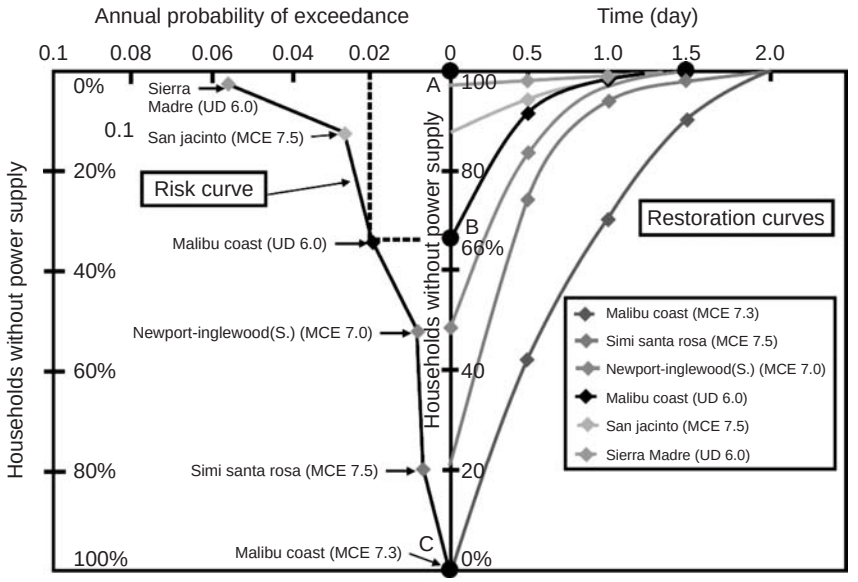
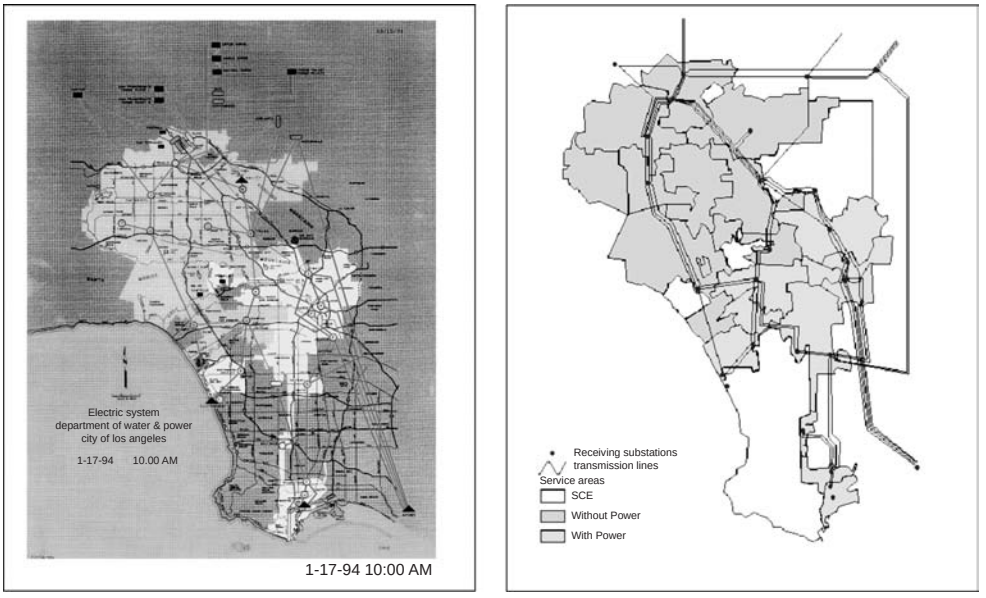


Figure 19 Risk Curves and Restoration Curves for LADWP's System.



(a) Reported by LADW (b) Simulated

Figure 20 State of LADWP's Power Supply Restoration at 6 Hours after the Northridge Earthquake.

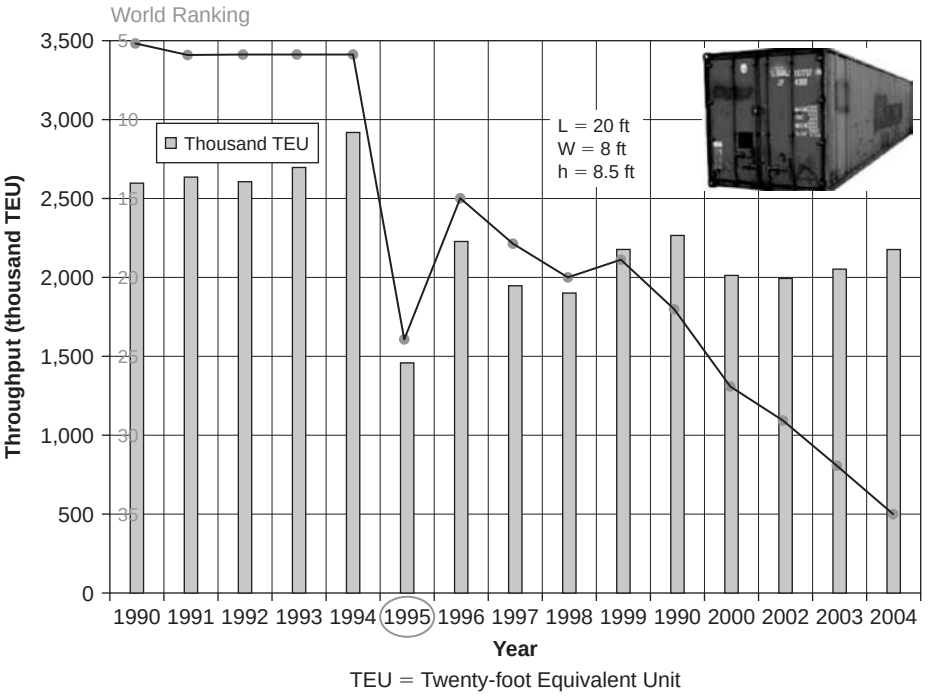


Figure 21 Yearly Throughput and World Ranking of Kobe Port.

7. Derived the likelihood function to estimate fragility parameters on the basis of multinomial damage model.
8. Developed simulation-based method for evaluation and actually evaluated, for the first time, the interaction between power and water systems due to power blackout.
9. Devised a probabilistic model for restoration process based on restorability of each component.

Acknowledgements

This work was supported by the Earthquake Engineering Research Centers Program of the National Science Foundation under Award Number EEC-9701471 through the Multidisciplinary Center for Earthquake Engineering Research.

References

Bruneau, M., Chang, S.E., Eguchi, R.T., Lee, G.C., O'Rourke, T.D., Reinhorn, A.M., Shinozuka, M., Tierney, K., Wallace, W.A. & von Winterfeldt, D., A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra*, Vol. 19, No. 4, pp. 733–752, 2003

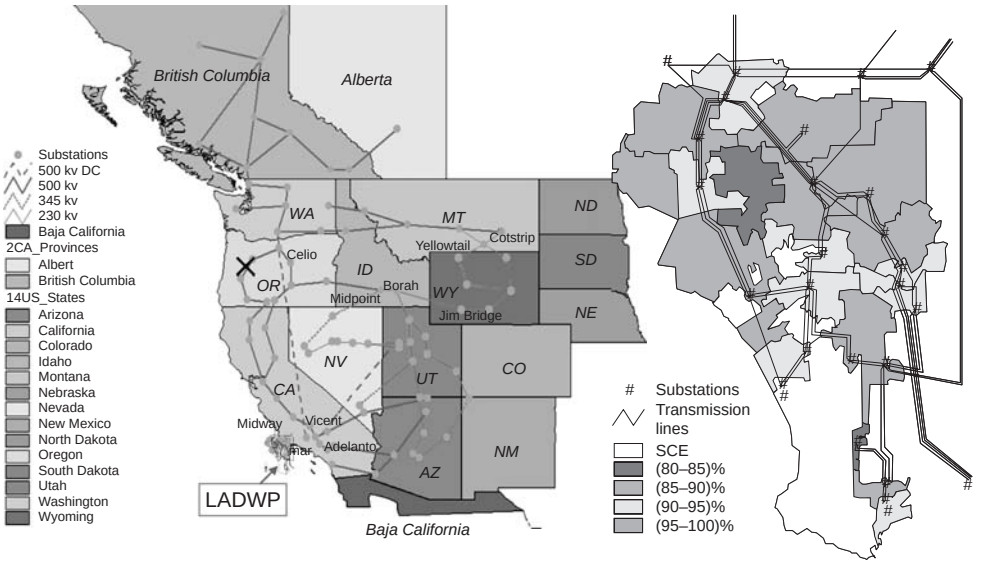


Figure 22(a) Impact of Remote Accident on LADWP.

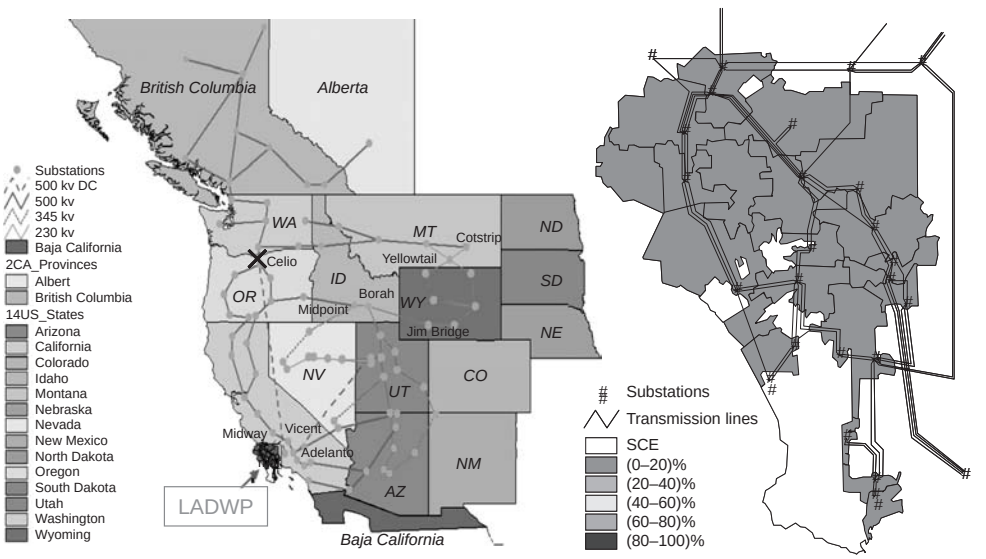


Figure 22(b) Impact of Remote Accident on LADWP.

Chang, S.E., Shinozuka, M. & Moore, J., Probabilistic Earthquake Scenarios: Extending Risk Analysis Methodologies to Spatially Distributed Systems. *Earthquake Spectra*, Vol. 16, No. 3, August, pp. 557–572, 2000

Chang, S.E., Svekla, W.D. & Shinozuka, M., Linking Infrastructure and Urban Economy: Simulation of Water Disruption Impacts in Earthquake. *Environment and Planning B*, Vol. 29, No. 2, pp. 281–301, 2002

Chang, S.E. & Seligson, H.A., Evaluating Mitigation of Urban Infrastructure Systems: Application to the Los Angeles Department of Water and Power. in *Proceedings of the 2003 Technical Council on Lifeline Earthquake Engineering Conference*, Reston, Virginia, 2003.

Dong, X.J., The Seismic Performance Analysis of Electric Power Systems. *PhD Dissertation*, University of Southern California, June 2002

Hall, J.F., (ed.) Northridge Earthquake of January 17, 1994: Reconnaissance Report. *Earthquake Spectra*, Supplement C to Vol.11, 1995

http://www.wecc.biz/main.html

Nojima, N. & Sugito, M., Development of a Probabilistic Assessment Model for Post-Earthquake Residual Capacity of Utility Lifeline Systems. J.E. Beavers, ed., *Advancing Mitigation Technologies and Disaster Response for Lifeline Systems*, Reston, Virginia: American Society of Civil Engineers, Technical Council on Lifeline Earthquake Engineering Monograph No.25, pp.707–716, 2003

Rose, A., Benavides, J., Chang, S.E., Szczesniak, P., & Lim, D., The Regional Economic Impact of an Earthquake: Direct and Indirect Effects of Electricity Lifeline Disruptions. *Journal of Regional Science*, Vol.37, No.3, pp.437–458, 1997

Saadeghvaziri, M.A. & Feng, M., Experimental and Analytical Study of Base-Isolation for Electric Power Equipment. *Research Progress and Accomplishments 2000-2001*, Multidisciplinary Center for Earthquake Engineering Research, May, pp. 29–40, 2001

Shinozuka, M., Cheng, T.C., Jin, X., Dong, X. & Penn, D., System Performance Analysis of Power Networks. *Proceedings of the Seventh U.S. National Conference on Earthquake Engineering* (7NCEE), Boston, Massachusetts, July 21–25, 2002

Shinozuka, M. & Dong, X., Seismic Performance Criteria for Lifeline Systems. *Proceedings of the Eighth U.S. – Japan Workshop on Earthquake Resistant Design of Lifeline Facilities and Countermeasures Against Soil Liquefaction*, Tokyo, Japan, Dec. 15–18, 2002.

Shinozuka, M. & Chang, S.E., Evaluating the Disaster Resilience of Power Networks and Grids. *Modeling Spatial Economic Impacts of Disasters*, Springer-Verlag Edited by Y. Okuyama and S.E. Chang, 2004

Tanaka, S., Shinozuka, M., Schiff, A. & Kawata, Y., Lifeline Seismic Performance of Electric Power Systems during the Northridge Earthquake. *Proceedings of the Northridge Earthquake Research Conference*, Los Angeles, California, August 20–22, 1997

U.S.-Canada Power System Outage Task Force (2003), Interim Report: Causes of the August 14th Blackout in the United States and Canada, ftp://www.nerc.com/pub/sys/all_updl/docs/blackout/814BlackoutReport.pdf

Webb, G.R., Tierney, K., & Dahlhamer, J.M., Businesses and Disasters: Empirical Patters and Unanswered Questions. *Natural Hazards Reviews*, Vol. 1, No. 2, pp. 83–90, 2000

Chapter 13

Application of reliability in earthquake engineering

Yi-Kwei Wen

Professor Emeritus of Civil Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA

ABSTRACT: Treatment of uncertainty is a top concern in performance evaluation and design of civil infrastructural systems against earthquakes because of the large uncertainty in both the excitation and the system capacity and potential large losses in the event of a damaging earthquake. Modern reliability concept and methodology have matured to the extent that proper treatment of these uncertainties can be included in the evaluation and design process such that engineers would have confidence in the performance of the system satisfying specific design objectives. Some of the recent developments and applications to design of structural systems are briefly reviewed and simple examples of application are given to demonstrate the methodology and procedure.

1 Introduction

Among the loadings on structures that engineers have to consider in performance evaluation and design, seismic loads are the most challenging due to the large uncertainty associated with the forces and the structural responses that it produces. The earthquake occurrence time, magnitude, rupture surface features, seismic wave attenuation and amplification, and finally the dynamic response behavior of the structure and the structural capacity against damage and collapse are subjects that cannot be predicted with certainty. To include these uncertainties and their effects in the structural performance evaluation and design, methods of probability and statistics are required.

Until recently, probabilistic treatment of seismic loads has been limited to the selection of design earthquake ground motion parameters based on return period such as peak or spectral acceleration. These ground motion parameters are then multiplied by a series of factors to arrive at the design seismic loads. The uncertainty is treated implicitly by allowing conservatism in the design forces based on professional judgment and experience and are calibrated such that the resultant designs do not deviate significantly from the acceptable practice at the time. As a result, the reliability of such design against damage and collapse is unknown and undefined.

The large losses suffered during recent earthquakes such as 1994 Northridge earthquake, 1995 Kobe earthquake, 1999 Turkey earthquake, and 1999 Chi-Chi earthquake, however, have prompted the profession to re-evaluate the whole design process and concentrate more on the uncertainty issue. As a result, large research efforts have been undertaken and significant progresses have been made in uncertainty modeling and applications to performance evaluation and design. The devastating earthquakes in Sichuan, China last May further accentuates the graveness and urgency

of this problem. Some of the more important developments are summarized in this paper.

Characterization of uncertainties as inherent variability (aleatory) versus modeling errors (epistemic) and their probabilistic treatment are first introduced. The uncertainties in earthquake engineering problems are then described in terms of demand on and capacity of the system. On the demand side, the uncertainty propagation from seismic source to path, site, ground motions, excitation intensity measures, structural responses, and system limit states. On the capacity side, uncertainty in material properties, member and system capacity against various limit states including incipient collapse are briefly described. The reliability analysis under seismic loads over a given period of time via the proposed FEMA/SAC procedure in US, then follow with numerical examples. Applications of the reliability analysis to probabilistic codes and standards are described.

2 Characterization of uncertainty

To understand the impact of uncertainty in earthquake engineering, it is instructive to first characterize uncertainty. Uncertainty associated with inherent variability is irreducible either because of the nature of the problem or our inability to reduce the uncertainty. The modeling errors in principle can be reduced as our knowledge accumulates. In the case of earthquake engineering, for example, the occurrence time, magnitude and distance of the next severe earthquake are inherently random. Whether a certain probability distribution and distribution parameters (mean and coefficient of variation), or an attenuation equation are correct based on a small number of observations of past events is part of the modeling errors. The same is true with structural capacity. The variability in steel yielding strength, member stiffness, damping ratio, and ductility capacity are random; the structural analysis models used in describing these parameters have modeling errors. As our knowledge about seismic events and structures improve with more observations, analyses, and experiments, better models with smaller uncertainty can be developed. In earthquake engineering, the emphasis so far has been more on the randomness; nevertheless the modeling errors and its impacts in performance evaluation and design begin to receive serious attention recently. From this point on, for simplicity, the inherent variability will be referred to as "randomness" and the modeling errors as "uncertainty".

Treatment of both randomness and uncertainty can be handled by probabilistic methods. The treatment of the randomness is generally modeled by a random variable if it does not change significantly with time. For example, in many situations we assume that the structural capacity can be modeled by a random variable. The random variable is then characterized by its moments (mean, standard deviation, skewness, etc.), and a probability distribution based on data or engineering judgment and experience. If it does change with time such as earthquake occurrence and ground acceleration during each occurrence, it can be treated by a random process such as a Poisson or Gaussian process. The parameters of these processes are estimated based on again data and engineering judgment and experience. Given the model and parameters, these random variable and random process models allow one to evaluate probability of event of interest to the engineer such as that of a prescribed structural limit state being exceeded over a given period of time, which in turn allows the engineer to make rational design

or retrofit decisions. The uncertainty in the models, however, also affects engineer's decision since a change in the model parameter values or model itself would definitely change the probability of structural limit state.

Treatment of uncertainty can be handled by the method of confidence interval. For example, because of small number of observations there is uncertainty associated with the mean occurrence rate of a Poisson process. Similarly, a small number of tests contribute to the uncertainty in the mean yield strength of a structural steel. The mean value calculated from the samples is also treated as a random variable with a standard deviation that decreases proportionally to the square root of the sample size. One can put bounds on this parameter, which depends on the confidence level associated with the bounds and the sample size. Alternatively, one can work with a random variable, which is the difference between the (unknown) true mean and sample mean and depends on the sample size and therefore account for the uncertainty. In structural reliability analysis, one can account for the uncertainty based on judgment and experience by allowing either excitation and resistance model parameters to be random variables and as a result the probability of limit state calculated will also be uncertain and can be described by confidence interval.

Consider a simple example of a structure with a random capacity X modeled by a normal distribution with mean μ_X and standard deviation σ_X under a deterministic demand d. The limit state probability accounting for the randomness only (with known μ_X and σ_X) is given by

$$P(X < d) = \Phi\left(\frac{d - \mu_X}{\sigma_X}\right) \tag{1}$$

in which Φ = cumulative distribution function of a standard normal variate. Now assume that σ_X is known, but μ_X is uncertain and modeled by a normal random variable with a mean $\overline{X}$ and a standard deviation S_X. The limit state probability given in Eq.1 is therefore also a random quantity with a probability distribution. One can show [Ang and Tang, 1984] that the mean value and q percentile value of P are given by

$$E[P] = \int_{-\infty}^{\infty} \Phi\left\{\frac{d - \mu_X}{\sigma_X}\right\} f_{\mu_X}(\mu)d\mu$$

$$= \Phi\left\{\frac{d - \overline{X}}{\sqrt{\sigma_x^2 + S_x^2}}\right\} \tag{2}$$

$$P_q = \Phi\left\{\frac{(d - \overline{X}) + S_x\Phi^{-1}(q)}{\sigma_X}\right\} \tag{3}$$

Note that the uncertainty measure S_x appears in both equations. Comparing Eq. 1 and Eq. 2, one can see that the net effect of additional uncertainty in μ_X is that it increases the total uncertainty and as result increases the limit state probability. From Eq. 3 one can calculate the confidence interval for the limit state probability to account for the epistemic uncertainty. Table 1 shows the dependence of confidence intervals

Table 1 Mean and percentile value of limit state probability as function of S_x/σ_X

($a = 1, \overline{X} = 2$, and $\sigma_X = 0.5$).

Percentile, q	$S_x/\sigma_X = 0.5$	$S_x/\sigma_X = 1.0$
5%	0.0029	0.00013
15%	0.0058	0.0012
25%	0.0097	0.0037
50%	0.023	0.023
mean	0.037	0.078
75%	0.048	0.093
85%	0.069	0.168
95%	0.119	0.361

on the epistemic uncertainty parameter S_x. From the Table 1 and Eq. 3 it is seen that the 50 percentile value of the limit state probability of 0.023 is equivalent to the case of assuming $\mu_X = \overline{X}$, since $\Phi^{-1}(0.5) = 0$ and Eq. 3 reduces to Eq.1 , i.e., the case in which only the randomness is considered. Note that the mean value estimate of the limit state probability is always higher than the median since the uncertainty in μ_X has been considered in the integral for the mean estimate. When the uncertainty in the mean capacity is half of the capacity randomness, one can use the 85% value of 0.069 to indicate that we are uncertain about the true limit state probability but are 85% confident that it is less than 0.069. Alternatively one can use the 15% to 85% interval of (0.0058 to 0.069) to indicate that due to the uncertainty, we are 70% confident that the true limit state probability will lie somewhere within the above interval. The interval increases to (0.0012 to 0.168) when the uncertainty is equal to the randomness ($S_X = \sigma_X$). Such wide intervals clearly indicate the importance of the effect of parameter uncertainty and modeling errors in the reliability performance evaluation. Often the nature of the problem and engineering convenience dictate the approach such as shown in the above example.

Another convenient method is treating the uncertainty as a multiplying factor applied to the randomness as follows [Ang and Tang, 1984].

$$X = N\tilde{X} \tag{4}$$

in which $\tilde{X}$ is the randomness and N is the correction factor to account for the uncertainty, also a random variable. μ_N = mean (bias) and δ_N = coefficient of variation (uncertainty) of N. The mean of X is then approximated by

$$\mu_X \approx \mu_N \, \mu_{\tilde{X}} \tag{5}$$

and the coefficient of variation X by

$$\delta_X \approx \sqrt{\delta_N^2 + \delta_{\tilde{X}}^2} \tag{6}$$

in which both randomness and uncertainty are considered and combined. Note that the above method can be extended to the situation when $\widetilde{X}$ is a function of set of random variables $\widetilde{X} = g(Y_1, Y_2, \ldots Y_n)$. Using a first order Taylor Series approximation for g, one can obtain

$$\mu_X \approx \mu_N \, \mu_g \tag{7}$$

in which,

$$\mu_g \approx g(\mu_{Y_1}, \mu_{Y_2}, \ldots \mu_{Y_n}) \tag{8}$$

and

$$\delta_X \approx \sqrt{\delta_N^2 + \delta_g^2} \tag{9}$$

in which

$$\delta_g \approx \frac{1}{\mu_g} \sum_i \sum_j \frac{\partial g}{\partial y_i}\bigg|_{\mu_{\underline{y}}} \frac{\partial g}{\partial y_j}\bigg|_{\mu_{\underline{y}}} \rho_{ij} \sigma_{Y_i} \sigma_{Y_j} \tag{10}$$

The partial derivatives are evaluated at the mean values of Y_i, $\mu_{\underline{y}}$; and ρ_{ij} denotes the correlation coefficient between Y_i and Y_j. $\rho_{ij} = 1$, when $i = j$, $\rho_{ij} = 0$ if Y_i and Y_j are statistically independent.

Uncertainty can be also handled by Bayesian method which approaches the problem from a different perspective and give different interpretation but yields results which are similar to the above approximate classical method.

3 Uncertainties in demand

In structural performance evaluation, it is convenient to describe the system performance in terms of demand and capacity. The demand can be the force (shear, bending moment, axial forces, overturning moment) or the response (displacement, velocity, acceleration, drift, ductility, energy dissipation) in the system caused by the ground excitation. The capacity of the system is the maximum forces or response that the system can withstand without member or system failure. The member or system failure in turn can be described by various limit states of interest to the engineers. For example, commonly used limit states in terms of response demand are drift limits corresponding to various performance requirements from immediate occupancy to collapse prevention. In theory, both the capacity and demand depend on the excitation and the structural property. Consider the performance of a structural system for a given period of time. The demand as described above such as system global drift is clearly a quantity that fluctuates in time and highly uncertain depending on the seismic excitation during the period. The capacity is primarily a property of the system which also depends on the type of excitation. It is common practice to use the maximum response or force over a given time period (annual or per 50 years) as the demand variable. For example, a 10% probability of exceedance in 50 years is commonly used as the probability threshold for selecting the design earthquake. The uncertainty in the demand so defined can be traced back to the chain of events that cause the response

or force as briefly described as follows. Details of course can be found in the extensive literature on these subjects.

3.1 *Earthquake source*

Seismic excitations come from events at different times and of different magnitudes, distances, focal depths and rupture surface geometries and features.

3.1.1 *Occurrence time*

The random occurrence in time can be modeled by random processes, such as from the simple Poisson process, to more involved renewal and Markov processes. These models allow one to calculate the probability of number of occurrences over a given period. The Poisson process occurrence rate is time-independent. The only parameter in the model is the annual rate of occurrence, ν. In spite of the rather restrictive assumption, it has been widely used. Based on the Poisson model, the number of occurrence, N, over an interval (0,t) is given by

$$P_t(N = r) = \frac{(\nu t)^r}{r!} e^{-\nu t} \tag{11}$$

The mean waiting time till the next occurrence is $1/\nu$, the return period.

In general, the Poisson model works well when there are a number of sources. Individually their occurrences may show time dependence but collectively the occurrences over time tend to become less dependent and can be approximated by a Poisson process. When dealing with individual and particularly large events such as characteristic earthquakes along well-defined faults with good records of the past occurrences, the latter two models are often used to incorporate the time dependence. For example, a renewal process has been used in the probabilistic prediction of large events in the western United States [Working Group, 1995] in which the random time intervals between occurrences are assumed to be a lognormal random variable. The probability of number of occurrences of a renewal process for a given period would depend on past history and the mathematics is more involved. Often the interest is on the next occurrence within t years from now knowing the last event occurred t_0 years before. Counting the time from the last event, the probability is given by

$$P(T < t_0 + t | T > t_0) = P(t_0 < T < t_0 + t)/P(T > t_0) \tag{12}$$

in which T is the inter-occurrence time, e.g., a lognormal random variable. The above probability can be easily calculated when the mean and standard deviation of T are known since the two parameters required in lognormal probability calculation in Eq. 12 can be determined from these two moments. Note that the probability depends not only on the time interval (0,t) but also on past occurrence time history given by t_0. Note also that the Poisson model would yield

$$P(T < t) = P_t(N \geq 1) = 1 - e^{-\nu t} \tag{13}$$

since the independence assumption implies that past history has no bearing on future occurrences. The difference between the two models could be significant. For example,

consider the characteristic event of 1964 Alaska earthquake. Assume an average recurrence time of 700 years and a coefficient of variation of 60%. According to Eq. 12 using a lognormal distribution for T with $t_0 = 44$ and $t = 50$, the probability of a repeat of the event in the period from 2008 to 2058 is 0.00028. According to Eq. 13 with $\nu = 1/700$ and $t = 50$, the probability is 0.069. The renewal model result is lower by a factor of 250, showing the effect of the dependency.

3.1.2 *Epicenter location*

The exact location of future earthquake epicenter is unknown. Random spatial distribution models can be used for this purpose. For example, line and areal source models where the epicenter is assumed to follow certain distribution on the line or within a well-defined region have been used for this purpose in the past. Such distributions can be obtained based on occurrence statistics collected from past earthquakes. For example, in the context of a Poisson occurrence model, one can express the mean occurrence rate of future events per unit area as function of the location $\nu(x,y)$ for an areal source and $\nu(\ell)$ as function along the line source. One can then evaluate the probability of occurrence of various events within the area or along the line. The occurrence rate of events in a given region, the random magnitude, and spatial distribution of epicenter given the occurrence in time can be used to model the temporal and spatial randomness of future events.

3.1.3 *Magnitude*

The magnitude variability is generally described by the Gutenberg-Richter equation, which expresses the logarithmic frequency as a linear equation of magnitude for certain range of the magnitude

$$\begin{aligned} &\mathrm{LogN} = a - bM \\ &\text{for } m_L < M < m_U \end{aligned} \tag{14}$$

The implication is that the variability of magnitude given the occurrence of an earthquake can be modeled by a truncated exponential probability density of the following form

$$\begin{aligned} &f_M(m) = Cbe^{-b(m-m_L)} \\ &\text{for } m_L < M < m_U \end{aligned} \tag{15}$$

The randomness in magnitude is therefore captured by the above distribution. Depending on the range, the variability in magnitude described by the above distribution in terms of coefficient of variation is quite large and close to one. When data are limited the uncertainty in parameters a, b, m_U, and m_L could be also important.

3.1.4 *Rupture surface*

There are many other random parameters of the source such as the size and geometry of the rupture surface, stress drop, and slip variation within the surface which could be

also important factors for consideration. The effects of the randomness of these parameters are to certain extent absorbed in the attenuation equation and seldom explicitly considered in seismic risk analysis. The exception is in simulation of individual large events. For example, random field models have been developed for the slip variation within the rupture surface and used in simulation of ground motions [e.g., Somerville *et al* 1997, Wen and Wu 2001].

3.1.5 *Path and site*

As the seismic waves propagate from the source through the rock and soil media to the ground surface at the site, they are attenuated or amplified and many factors contribute to the uncertainty in the attenuation and amplification processes. As mentioned above, the effects of many other random parameters associated with the source are also included in the attenuation model. As a result, the randomness in the attenuation model is usually very large as can be seen from the large scatter of attenuation of various ground motion intensity measures such as spectral ground acceleration and velocity based on observations during past earthquakes. As shown in previous chapters, the forms of the attenuation equations are usually a result of wave propagation theory modified by observational results. The most important independent variables in the attenuation equations are the magnitude (M), distance (R), and site soil classification (S). In view of the large uncertainty, the attenuation equation A(M, R, S) generally describes the central value and the scatter is modeled by a random variable. When the intensity measures are plotted on a logarithmic graph, the scatter generally follows approximately a normal distribution. Therefore given M, R, and S, the intensity measure, e.g. spectral acceleration S_a at the site is approximately a lognormal random variable with expected (mean) value E[log S_a] described by the attenuation equation; i.e.

$$\mathrm{E}[\log \mathrm{S_a}(\mathrm{M}, \mathrm{R}, \mathrm{S})] = \mathrm{A}(\mathrm{M}, \mathrm{R}, \mathrm{S}) \tag{16}$$

The scatter is given by $\sigma_{\log S_a}$, in which σ denotes standard deviation. σ in general is also a function of M and S [e.g. Boore and Joyner, 1994] but usually is regarded as a constant as an approximation. Therefore, in such a formulation, all the randomness in wave propagation from the source to the site and some randomness associated with the source is captured by $\sigma_{\log S_a}$. Note that the mean and standard deviation in these equations are in terms of log S_a not S_a. After proper conversion, the mean and coefficient of variation of S_a can be shown to be

$$\mathrm{E}(\mathrm{S_a}) = \exp[2.3\mathrm{A} + 0.5(2.3\sigma)^2] \tag{17}$$

$$\delta_{S_a} = \sqrt{e^{(2.3\sigma)^2} - 1} \tag{18}$$

in which A and σ are the attenuation equation prediction and scatter in log scale. For example, a scatter of $\sigma_{\log S_a} = 0.3$, a value commonly seen in attenuation equations,

actually means a coefficient of variation of 78% in S_a. The probability that S_a exceeds a given limit of a_0 is therefore given by

$$P(S_a > a_0|M, R, S) = 1 - \Phi\left[\frac{\ln a_0 - 2.3A(M, R, S)}{2.3\sigma}\right] \tag{19}$$

Note that the above equation describes the randomness in attenuation alone when M, R, and S are known. M, R, and S are also random variables, which would influence the demand on the structural system. Also, the uncertainty in attenuation equation itself (modeling errors) is evident from the various forms of attenuation equations for the same region, which give different results. Again such uncertainty is not generally considered explicitly.

3.2 *Ground excitation and structural response*

The demand on the structure over a given time period in the future is the ground motions and structural responses that they produce. They are unpredictable and random functions of time. In theory, they can be modeled by a continuous random process whose parameters depend on the source, path, site, and structural characteristics. The ground excitation given the occurrence of an earthquake in the future is therefore a continuous random process of time that depends on magnitude m, distance r, and site condition, i.e. a(tlm, r, s). The structural response in turn is also a random process depending on the excitation and the structural characteristics and the excitation parameters. Although such random process models have been developed for both the excitation and structural responses based on random process theory and method of random vibration [e.g. Wen 1989, 1990], the nonstationarity in the excitation and quite often nonlinear and inelastic dynamic response of the system render the theoretical treatment difficult for real structural systems.

3.2.1 *Excitation intensity measures*

In performance evaluation, the structural response demands are often described in terms of the maximum responses such as maximum global displacement, interstory drift or energy dissipation over the duration of the excitation. These demand variables are random and the annual maximum or maximum value over 50 years is customarily used. The uncertainty in these demand variables can be traced back to those in the structural characteristics as well as source, path, and site parameters. The propagation of uncertainty along the chain of events that lead to the demand variable is a rather complicated process involving random variables and random processes and linear and nonlinear input-output relationship. To simplify the problem, engineers have been trying to find some key ground excitation intensity measures that correlate well with the structural demand variable. The peak ground acceleration, velocity, and displacement have been traditionally used for this purpose. These measures generally show poor correlation with the structural response since the structural characteristics are not considered.

Luco and Cornell [2002] examined a number of intensity measures that reflect the structural characteristics such as fundamental period and damping ratio based on

extensive regression analyses of steel structural systems of different design and configurations under excitation of recorded ground motions. The results showed that the spectral acceleration or displacement at the structure's fundamental period corresponding to a damping ratio of 5% generally predicts structural response well. To incorporate effects of higher modes and inelastic response, intensity measures consist of the combined first and second mode spectral acceleration or displacement and first mode elastic and inelastic spectral accelerations were examined. They give even better results as indicated by the smaller scatter in the regression relationship compared with using only the fundamental period elastic response. This is achieved, however, at the expense of more complicated form of the intensity measure. One advantage of using spectral response variable is that these quantities can be related directly to M, R, and S via the attenuation equation and additional dependence of structural response on M and R are small and can be ignored in approximation [Shome et al, 1998]. To consider the effect of bi-axial excitation, Wang and Wen [2000] also proposed a bi-directional spectral displacement defined as the maximum of the vector sum of the displacements in two principal directions at the two fundamental periods of the structure in the two principal directions. It can be used to better correlate with the bi-axial structural response measure such as bi-axial drift ratio, which is defined in the same way.

3.3 *Seismic hazard analysis*

The uncertainty in the seismic excitation can be therefore approximately described in terms of a random variable of the above intensity measure such as the maximum spectral acceleration over a given period of one year or fifty years. The probability of exceedance of such a random variable is generally referred to as the seismic hazard curve. For example, if the spectral acceleration S_a is used, the probability of exceedance in t (e.g. 50) years is given by

$$P_t(S_a > a) = H_t(a) \tag{20}$$

$H_t(a)$ is the hazard curve, which can be constructed from the probabilistic models of the source, path, and site as described above based on available regional seismicity information. For example, consider a region in which there is a well-defined fault of characteristic earthquakes of known magnitude. The probabilistic distribution of the inter-occurrence time and date of last occurrence are also known. There is also an areal source of smaller events whose occurrences can be modeled by a Poisson process with an occurrence rate which is a function of the location, $\nu(x, y)$; and whose magnitude can be modeled by an exponential distribution based on a Gutenberg-Richter equation. In addition, there is also a line source along which the occurrence rate $\nu(\ell)$ and magnitude distribution of future events are known. The attenuation equations for events from these sources have also been established. Assuming the events from these three sources are statistically independent, the seismic hazard over the next t years can be evaluated as follows:

$$P_t(S_a > a) = 1 - [P_C(S_a < a|C)P(C)][P_A(S_a < a)][P_L(S_a < a)] \tag{21}$$

In which C denotes occurrence of characteristic events modeled by Eq. 12; the conditional probability of spectral acceleration given the occurrence of the event can be

estimated from the attenuation equation models. A and L refer to the areal and line sources. The last two terms in Eq. 21 are obtained by considering contribution from all future events within the areal and line sources and the occurrence as a Poisson process as follows

$$P_A(S_a < a) = e^{-t \int_x \int_y \int_m \nu(x,y) P(S_a > a|m,r,s) f_{MA}(m) dx dy dm} \tag{22}$$

$$P_L(S_a < a) = e^{-t \int_\ell \int_m \nu(\ell) P(S_a > a|m,r,s) f_{ML(m)} d\ell dm} \tag{23}$$

in which subscript MA and ML refer to magnitude of events in the areal and line sources. The above procedure allows one to evaluate the spectral acceleration of different periods corresponding to a given probability of exceedance. The resulting response spectra are called uniform-hazard spectra (UHRS). The commonly used probability of exceedance is 50%, 10%, and 2% in 50 years such as in the USGS National Earthquake Hazard Mapping Project [Frankel *et al*, 1996]. The UHRS therefore is an efficient way of describing the seismic hazard and ground motion demand on the structure since the response of a linear structure corresponding to the above probability of exceedance can be easily predicted using the well-known modal superposition method.

For nonlinear systems, the UHRS cannot be directly used since modal superposition method can no longer be applied. There have been large efforts in the past on extension of the concept of UHRS to nonlinear inelastic systems. Based on investigation of large number of single-degree-of-freedom (SDOF) systems under recorded ground motions, uniform-hazard inelastic response spectra have been established by researchers [Nassar and Krawinkler 1992, Miranda and Bertero 1994, Collins et al 1996]. Empirical rules have been developed so that the uniform-hazard inelastic response spectra (UHIRS) can be constructed from the linear UHRS. The spectra give the ductility ratio of a SDOF system of given period and yield strength corresponding to a given probability of exceedance. The UHIRS therefore describes the demand on an SDOF inelastic system. It is mentioned that most real structural systems cannot be adequately described by a SDOF system since the effect of higher modes cannot be included; hence the application of UHIRS is limited.

3.4 *Modeling of epistemic uncertainty by logic tree*

When dealing with uncertainty in the selection of magnitude, recurrence model, and attenuation equation, etc., in seismic hazard analysis, a logic tree is frequency used with branches for different models or values, each with assigned likelihood based on judgment/experience [e.g. Frankel et al 1996, Frankel 1997]. It is therefore a method for treating the epistemic uncertainty. At each branch of the tree, further characteristics and uncertainty can be assigned in accordance with the expert's opinion [e.g., SSHAC, 1995]. For example, referring to Fig. 1, going from the site to the source, possible attenuation equations are first identified. The occurrence model is either a memoryless Poisson process according to the Gutenburg-Richter equation, a renewal process for characteristic events with a specified recurrence time distribution, or a Markov process with memory specified by a transition matrix. At each branch, candidate models or equations are assigned with a relative likelihood reflecting the judgment

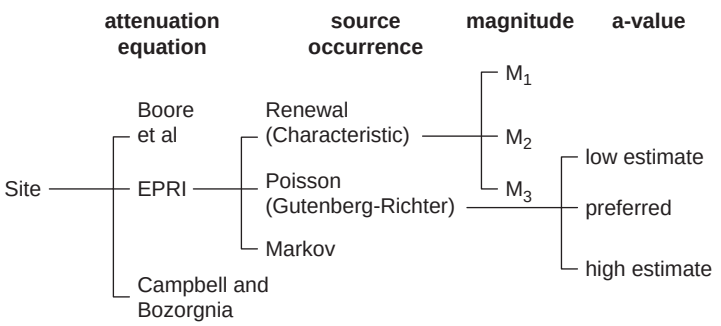


Figure 1 Logic tree in seismic hazard analysis.

and experience of the experts. For example, if it is a characteristic event, the possible choices of magnitude are M_1, M_2, or M_3 with given relative likelihood.

In the seismic hazard analysis these relative likelihood of the magnitude is then converted into discrete probability mass function and incorporated into the risk analysis. For example, because of the modeling uncertainty, the result of the seismic hazard analysis as given in Eq. 21 becomes a random variable. A common practice is to determine the expected (mean) value of the seismic risk by integration (or summation) over all possible combinations of these values weighted by their likelihood (or probability mass). A simpler and more convenient way is to approximate the mean risk estimate by using the mean value at each branch when possible (such as the a and b values in the Gutenburg-Richter equation or the magnitude of a characteristic earthquake) and reduce the number of possible combinations and hence the required numerical effort. The implications of using the mean value versus the percentile value will be illustrated in an example in the following sections.

3.5 *Probabilistic structural response demand analysis*

To establish the probabilistic relationship between the ground motion intensity measure and the response of MDOF nonlinear systems, one can use method of random vibration with the ground motion modeled as a random process. Alternatively one can use a regression analysis of nonlinear time history responses under recorded ground motions. Because of the inherent nonstationary nature of the random excitations in earthquakes and the analytical difficulty in modeling complex nonlinear member response behaviors such as brittle fracture in random vibration, the time-history/regression analysis is more practical. To cover a wide range of excitation and structural response in the inelastic range, such ground motions are frequency scaled. Although the frequency content and duration of ground motions due to events of different magnitudes and distances are different and scaling may violate the basic physics of ground motions, results based on extensive investigations [Shome *et al*, 1998] show that when intensity measure such as spectral acceleration is used in the scaling, the errors are small. The regression analyses therefore allow one to establish the functional relationship between the intensity measure and the structural demand variables such as global (roof) and

local (interstory) drifts and energy dissipation (cumulative damage). In the following, the relationship between maximum interstory drift and spectral acceleration is used as an example. The method can be applied to other structural demands under a different intensity measure such as spectral displacement or bi-directional spectral displacement [Wang and Wen, 2000].

Based on extensive regression analyses of response of steel structures, Cornell et al (2002) proposed that the maximum interstory drift can be expressed as a simple power function of the spectral acceleration:

$$D = a(S_a)^b \tag{24}$$

Such relationship is necessarily approximate and there are large scatter around the regression line, which will be incorporated in the performance analysis as shown in the following. The regression prediction is therefore the estimate of the mean demand conditional on a given value of the excitation intensity measure, $E(D|S_a = a)$. The scatter in terms of the coefficient of variation, $\beta_{D|S_a=a}$, also depends on the intensity but again often regarded as constant as an approximation. The structural response demand given the excitation intensity therefore can be described by a random variable of a given distribution. The lognormal distribution generally gives a good fit, which can be used to describe the randomness in structural demand variable due to ground motion record-to-record variation even though these ground motions are of the same S_a. The probability of the structural demand being exceeded in t years can therefore be evaluated by the total probability theorem to incorporate the contribution from all values of S_a.

$$P_t(D > d) = \int_0^\infty P(D > d|S_a = a)H_t'(a)da \tag{25}$$

in which

$$P(D > d|S_a = a) = 1 - \Phi\left(\frac{\ln d - \lambda}{\beta_{D|S_a=a}}\right)$$

$$\lambda = \ln[a(S_a)^b] - 0.5\beta^2_{D|S_a=a} = \ln \widetilde{D},\ \widetilde{D} = \text{median value of D}$$

$$H_t'(a) = \text{derivative of seismic hazard curve}$$

Note that the calculation as shown in the above general analytical procedure could be quite involved and has to be carried out numerically. In code procedures and for a fast and approximate evaluation, closed form solution is desirable. It has been shown [Cornell *et al*, 2002] that if the result of the seismic hazard analysis as given above can be approximately described by a power law

$$H_t(a) = k_0 a^{-k} \tag{26}$$

in which k specifies the hazard decay and k_0 is the scale factor. The above lognormal distribution assumption for the demand given excitation is valid, Eq. 25 can be evaluated in a closed form

$$P_t(D > d) = H_t(a^d)\exp\left[\frac{1}{2}\frac{k^2}{b^2}\beta^2_{D|S_a=a}\right] \tag{27}$$

in which a^d is the spectral acceleration level corresponding to the demand d according to Eq. 24. Eq. 27 therefore describes a probabilistic structural response demand curve in which all the important randomness in excitation and structural response is considered. The first term is the demand curve without consideration of the randomness in the response-excitation relationship. The exponent function is the correction for this randomness. Note that the correction factor involves both the structural (b) and hazard (k) parameters. Note that Eq.26 is intended for approximating the tail distribution. It is no longer valid when the spectral acceleration is very small. The method is demonstrated by a simple numerical example as follows.

Consider a 3-story steel structural building with a fundamental period of 1 sec at a location where the 50-year seismic hazard and maximum interstory drift ratio as function of the spectral acceleration can be described by

$$H_{50}(a) = 0.0068a^{-3} \tag{28}$$

$$D = 0.06a^{1.2} \tag{29}$$

The hazard is such that the spectral acceleration at 1 sec is 0.4 g corresponding to an exceedance probability of 10% in 50 years and 0.7 g corresponding to 2% in 50 years, typical values for a site in the Los Angeles area. The building response is such that the maximum interstory drift ratio is 2% at a spectral acceleration of 0.4 g and 4% at 0.7 g, reasonable values for such steel building. Assuming the randomness in the drift ratio-spectral acceleration regression analysis $\beta_{D|S_a} = 0.3$, the 50-year probabilistic maximum interstory drift demand curve is then according to Eq. 27,

$$P_{50}(D > d) = 0.0068\left[\left(\frac{d}{0.06}\right)^{\frac{1}{1.2}}\right]^{-3}\exp\left[0.5\left(\frac{3}{1.2}\right)^2 0.3^2\right] \tag{30}$$

The correction factor to account for the randomness in the demand as given in the exponential function is 1.33 and the 50-year demand curve is simplified to $0.009904(d/0.06)^{-2.5}$ and shown in Fig. 2.

As demonstrated in previous sections, uncertainties either statistical in nature such as sampling errors or empirical in nature based on judgment could be important and have not been accounted for in the above formulation. These include uncertainties in the structural response analysis methods, choice of probability distributions and parameters and assumptions and approximations used in the source, path, and site parameters. Treatment and impact of these uncertainties will be covered more in details in probabilistic performance evaluation.

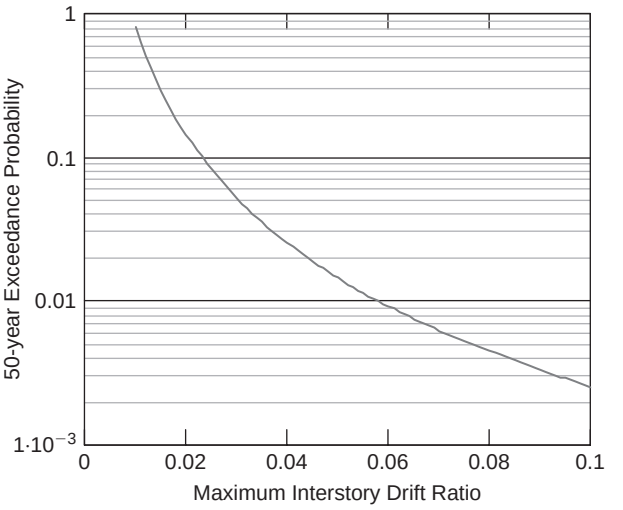


Figure 2 50-year maximum interstory drift ratio demand curve.

4 Uncertainties in capacity

Structural capacity, is defined as the maximum force, displacement, velocity, or acceleration that a member or a system can withstand without failure, or more specifically, without reaching a prescribed limit state. The capacity is therefore dependent on the material characteristics, member dimensions, system configuration, the limit state under consideration, and methods and models used in describing the capacity. As in the case of demand, both (aleatory) randomness and (epistemic) uncertainty are important elements in the evaluation of capacity and need to be carefully considered. In the following the capacity uncertainty and probabilistic treatments are described of first the construction materials, and then structural members and finally structural systems. Since capacity is always related to the limit state under consideration, some of the more frequently used limit states in performance evaluation will also be discussed. Again there is a large body of literature on this subject; emphasis here is on the uncertainty treatment.

4.1 *Material characteristics*

The member and system capacity depend directly on the material strength, which is inherently random. The randomness can be modeled by random variable based on test data. It is common to use the first two moments, i.e. the mean and standard deviation (or coefficient of variation), to describe the central value and the variability. Normal or lognormal distribution is commonly used for convenience. The actual strength of the material of a given member may differ significantly from the nominal values used in member capacity calculation. The correspondence between the nominal value and the actual value therefore needs to be established to estimate the real member capacity. The strength variability obviously depends on the material, manufacturing process,

and sometimes the testing protocol. In general, the variability in masonry and timber construction material is larger than those in reinforced concrete and steel. Material property variability and test data up to 1980 can be found in the report by Ellingwood *et al* [1980]. For example, the coefficient of variation of strength of timber varies in the range from 10% to 30% depending on species and in flexure or compression; and that of masonry walls from 10% to 26% depending on configuration and in compression or flexure. The coefficient of variation of compressive and tensile strength of concrete is around 18% and that of the yielding strength of steel reinforcement and steel member elements is around 10% or less. Characteristics of construction material such as concrete and structural steel evolve over time. Strength statistics of newer material such as high-strength steel and concrete may be found in more recent literature. For example, statistics on yield and ultimate strength of structural steel under various environmental conditions can be found in the recent FEMA/SAC report [2000].

4.2 *Uncertainty in member capacity*

4.2.1 *Member capacity under monotonic load*

The inherent randomness in the material property carries over to the structural members made of these construction materials. In addition, there is randomness in the dimensions of the members and also the capacity refers to a particular limit state such as shear, bending, or buckling failure under monotonic or cyclic loading condition. The randomness in terms of the bias (mean capacity/nominal capacity) and coefficient of variation of steel, reinforced concrete, masonry and glulam structural members (beams, columns, and walls) of various configurations and for various limit states can be found in Ellingwood *et al* [1982]. The majority of the bias factor is between 1.0 and 1.2 and the coefficient of variation under 20%. Normal or lognormal distribution has been used to model the capacity randomness. The difference between the two models is small when the coefficient of variation is small.

4.2.2 *Members capacity under cyclic load-damage index*

For seismic loading, one is especially interested in the member capacity under cyclic loading since members in a structural system generally undergo stress reversals of various amplitudes and the member may reach a limit state under combined action of large deflection and cumulative damage. To account for both effects, various damage indices have been proposed. The most widely used is the index developed by Park and Ang [1985] based on test results of 403 reinforced concrete members. The index is a linear function of maximum displacement δ_m and total hysteretic energy dissipation normalized by member ultimate displacement and monotonic loading δ_u, and yield force Q_y.

$$D = \frac{\delta_m}{\delta_u} + \frac{\beta}{Q_y \delta_u} \int dE \tag{31}$$

Different value of the index corresponds to different limit states such as 0.4 for serious damage and 1 corresponding to complete damage (collapse). Test data show that the damage index capacity of reinforced concreted member can be modeled by a lognormal

random variable with a mean value equal to 1.0 and a coefficient of variation of 0.53. It indicates that the randomness in the reinforced concrete member capacity is quite large. The index has been used in damage evaluation of buildings and other structures, e.g. Park *et al* [1985].

4.2.3 *Rotation capacity of steel connection members*

An important structural member in steel buildings is the connections between beams and columns. After the large number of brittle fracture failures found in many buildings due to 1994 Northridge earthquake, the capacity of connections against rotation demand under cyclic loading during earthquake excitations has attracted much attention of the structural profession. In the FEMA/SAC project [SAC, 2000], a comprehensive testing program of a large number (120) of welded and bolted connections of various configurations has been carried out according to pre-Northridge practice and for post-Northridge design in which many different improvements were incorporated. Test results of hundreds of experiments prior to 1994 were also analyzed. The connection rotation capacities for both pre- and post-Northridge connections were obtained. The capacity is defined in accordance with two limit states; θ_p the rotation limit when plastic deformation occurs and θ_g, the rotation limit corresponding to severe damage that the gravity load carrying capacity of the member is compromised. Test data generally show the dependence on the depth of the beam or the depth of the connection element of these capacities and large scatter. The mean values and standard deviations as linear functions of the depth of the beams were established from regression analyses of test results. Depending on the specific connection type and the depth of the beam, the rotation capacity and variability in terms of these two statistics shows large variation. For example, the capacity of the post-Northridge welded-flange-bolted-web connections has the following means and standard deviations,

$$\mu_{\theta_p} = 0.021 - 0.0003d_b \tag{32}$$

$$\sigma_{\theta_p} = 0.012 - 0.0004d_b$$

and

$$\mu_{\theta_g} = 0.050 - 0.0006d_b \tag{33}$$

$$\sigma_{\theta_g} = 0.011 + 0.0004d_b$$

For such a connection with a beam depth of 24 inches, $\mu_{\theta_p} = 0.0138$, and a standard deviation $\sigma_{\theta_p} = 0.0024$ (or a coefficient of variation $\delta_{\theta_p} = 17.4\%$); and $\mu_{\theta_g} = 0.0256$ and a standard deviation $\sigma_{\theta_g} = 0.00206(\delta_{\theta_g} = 8\%)$. The variability is moderate. For a free-flange and welded web connection with a beam depth of 36 inches, the regression results give $\mu_{\theta_p} = 0.0238$ and $\sigma_{\theta_p} = 0.0032(\delta_{\theta_p} = 134\%)$; and $\mu_{\theta_g} = 0.0364$ and $\sigma_{\theta_g} = 0.0604(\delta_{\theta_g} = 166\%)$. The variability is very large. Such large variation of coefficient of variation for different connections could partly due to the small number of samples used in the regression analysis. No distribution models were recommended for the capacity. In view of the small sample size and large coefficient of the variation, selection of the distribution model should be done with care. Note that with a

distribution model, say a normal distribution, one can predict the probability of limit state of plastic deformation or loss of gravity load carrying capacity of the connection member when the rotation demand θ_d is known

$$P(\text{plastic deformation}) = P(\theta_p < \theta_d) = \Phi\left(\frac{\theta_d - \mu_{\theta_p}}{\sigma_{\theta_p}}\right) \tag{34}$$

$$P(\text{loss of gravity load carry capacity}) = P(\theta_g < \theta_d) = \Phi\left(\frac{\theta_d - \mu_{\theta_g}}{\sigma_{\theta_g}}\right) \tag{35}$$

Probability of capacity being exceeded given the demand is also generally referred to as the fragility function. Since under earthquake excitations the demand is also a random variable, the limit state probability can be evaluated based on a reliability analysis as shown in the next section.

4.3 *Uncertainty in system capacity*

The description of uncertainty in system capacity is more involved since a structural system consists of many components and the system behavior is complex under dynamic excitation, especially when the system goes into nonlinear range. The system capacity can be therefore more conveniently described in terms of the system limit states of interest.

4.3.1 *System capacity against damage*

Commonly used system limit states are those corresponding to different damage states and performance levels. For example in SEOAC Vision 2000 [1995], there are five performance (damage) levels, fully operational (negligible), operational (light), life safety (moderate), near collapse(severe), and collapse(complete) and each level is related to a structural response level indicated by a transient and a permanent drift limit. In the FEMA/SAC project for steel buildings, the performance/damage levels were reduced to a more manageable two, immediate occupancy and collapse prevention. The system capacity is again described in terms of interstory drift angles. The uncertainty in the system capacity therefore can be described in terms of the drift capacity for different performance levels, such as the median drift capacity and its coefficient of variation. The commonly accepted distribution for the capacity is the lognormal distribution for its convenience in reliability analysis and reliability-based design as will be seen in the next section. Structures of different construction material, configurations, and designs would have different drift thresholds. Determination of drift capacities for different performance levels is largely a process of combination of analysis and judgment/experience. The determination of system collapse prevention capacity is discussed further in the following.

4.3.2 *System capacity against collapse-incremental dynamic analysis*

Of all the limit states and the corresponding system capacities, system collapse is the most difficult to determine. The reason is obvious that the structural dynamics close to collapse is extremely complex and is still largely an unsolved problem due to

nonlinear member and system response behaviors. The large record-to-record variation of ground motions and structural response behaviors further complicate the matter. Collapse of structures under random excitations is a difficult mathematical problem of stochastic stability. Solutions can be obtained only for simple idealized systems under excitations of simple stochastic processes such as white noise. Engineers have used an inelastic static pushover analysis in the past to estimate this capacity. It provides insight into the structural response behavior at large displacement but considers the first mode static response only, which is basically different from dynamic response. As a result, such analysis generally over-predicts the response and underestimates the capacity. Vamvatsikos and Cornell [2001] extended the concept of pushover analysis to dynamic response in the form of incremental dynamic analysis (IDA). The system capacity against collapse is evaluated by dynamic response analyses of the system under a suite of ground motion time histories such as the SAC ground motions. Each time history is scaled according to the spectral acceleration such that the structural response goes from linear elastic range to nonlinear inelastic and finally becomes unstable, i.e. a large increase in response with a small increase in the spectral acceleration. The displacement at the transition point is defined as the system displacement capacity against collapse. As mentioned earlier, due to the large record-to-record variation of the ground motions and extremely complex structural nonlinear behavior, the transition point is not always easy to pinpoint. Engineering judgments are often necessary and there are large scatters for different excitations with the same spectral acceleration. Fig. 3 shows an example of the interstory drift capacity using IDA of a steel moment frame under ground motions at Los Angeles recommended in the SAC project (Liao et al 2007). The square dots indicate the transition points. The uncertainty in capacity against collapse can be described in terms of the mean and standard deviation of the interstory drift capacity under multiple recorded ground motions from IDA. The coefficient of variation of this displacement capacity is generally of the order of 30%. Such a procedure has been used in the FEMA/SAC procedure.

5 Reliability of structrual systems

In view of the large uncertainties in both demand and capacity as shown in the previous sections, the performance of the structural systems can be described meaningfully only when these uncertainties are taken into consideration explicitly. In other words, evaluation of the performance needs to be described in terms of reliability of the structural system against various limit states over a given period of time. The structural performance is described by a set of random variables representing the uncertainty in the problems. The reliability problem is then solved using the first two moments. They are commonly referred to as the first order reliability method (FORM) or the second order reliability method (SORM). In earthquake engineering, an even simpler formulation of the problem is used in terms of two variables, demand versus capacity, as described in the following sections, for a given limit state.

5.1 *Demand versus capacity formulation in FEMA/SAC procedure*

The reliability problem is simplified considerably if the limit state can be stated in terms of the demand exceeding the capacity. Although it may be an over-simplification in that the capacity and demand may not always easily defined for certain limit states such as

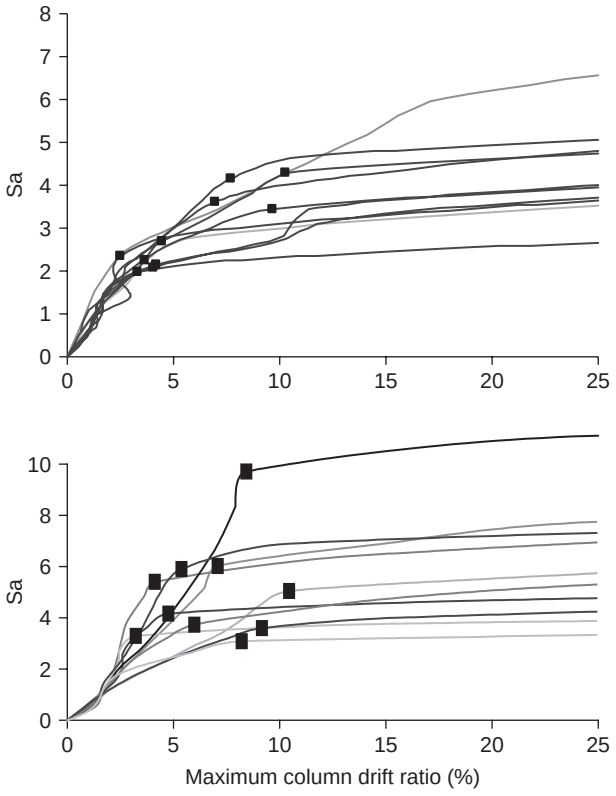


Figure 3 Example of IDA analysis of pre-Northridge (upper) and post-Northridge (lower) steel moment frame buildings (Liao et al 2007).

the case of system collapse discussed in the previous section. In earthquake engineering applications, simplicity nevertheless offers some advantages, and especially in code procedure formulation. This is the method used in the reliability based, performance oriented design procedure proposed in the SAC/FEMA Steel Project [Cornell *et al*, 2002], which is described in the following.

Considering now the limit state described in terms of only two random variables; R (capacity) and S (demand), the probability of limit state over a given period of time, t, is then given by the probability integral

$$P_t = P_t(R < S) = \int_0^\infty P_t(R \leq S|S = s) f_S(s) ds \tag{36}$$

Simple closed form solutions of the integration can be obtained when both R and S can be modeled by normal or lognormal random variables.

$$P_t = 1 - \Phi(\theta) \tag{37}$$

$$\theta = \frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2}}, \text{ when both R and S are normal}$$

$$= \frac{\lambda_R - \lambda_S}{\sqrt{\beta_R^2 + \beta_S^2}}, \text{ when both R and S are lognormal}$$

in which, $\beta_x = \sqrt{\ln(1 + \delta_x^2)}$, and $\lambda = \ln \mu_x - 0.5\beta_x^2 = \ln \tilde{x}$. $\tilde{x}$= median value of X

This is not the case when the demand described by Eq. 25 is not a simple normal or lognormal variable. The closed form solution, however, can still be obtained when the capacity variable R can be modeled by a lognormal random variable and the seismic hazard can be described by a power function as given in Eq. 26.

5.2 *Limit state probability considering randomness only*

Assume that the capacity randomness can be modeled with a lognormal variate with a median value $\tilde{C}$ and dispersion coefficient β_{CR}. It can be shown that the limit state probability is given by

$$P_t = H_t(a^{\tilde{C}}) \exp\left[\frac{1}{2}\frac{k^2}{b^2}(\beta_{D|S_a=a}^2 + \beta_{CR}^2)\right] \tag{38}$$

Note that the limit state probability consists of the probability of the spectral acceleration exceeding the median structural capacity multiplied by a correction factor depending on the randomness in the demand and the capacity, as well as the hazard characteristics (k) and demand/capacity relationship (b). The limit state probability as given in Eq.38 therefore accounts for all randomness in the problem.

Continuing with the example problem in 7.3.3, and assuming that the median drift ratio capacity of this 3-story building against incipient collapse is 0.05 with a dispersion parameter $\beta_{CR} = 0.35$, the 50-year incipient collapse probability of the building is given by

$$P_{50} = H_{50}(a^{\tilde{C}}) \exp\left[\frac{1}{2}\frac{3^2}{1.2^2}(0.30^2 + 0.35^2\right]$$

$$= H_{50}(0.86)e^{0.664} = 0.0107 \times 1.94 = 0.0208 \tag{39}$$

Which corresponds to an annual probability of 0.42×10^{-3}, or a return period of 2380 years

5.3 *Impact of Uncertainty*

Referring to the simple example problem on the impact of (epistemic) uncertainty on the limit state probability in Section 2 (Eqs. 1 to 3), one can see that when the uncertainty is considered the limit state probability becomes a random variable and needs to be treated as such. Depending on the application, for example, one may want to evaluate the mean value of the limit state probability to perform a risk/benefit analysis or evaluate the percentile value for a confidence interval estimate. The uncertainty can

be conveniently grouped into those in the hazard analysis, excitation/demand relationship, and structural capacity estimate. For example, the parameters k_0 and k in the seismic hazard model (Eq. 26), a and b in the regression equation for structural response (Eq. 24), and the parameters in structural capacity models (Eqs. 32 to 35) may all have uncertainty due to either modeling (e.g. incorrect functional form) or sampling (small number of test results) errors. For simplicity and tractability in analysis, the uncertainties in the seismic hazard, structural demand, and structural capacity models are assumed to be lognormal variables with a median values given by the model predictions and dispersion parameters β_{HU}, β_{DU}, and β_{CU}. The subscript H, D, and C denote hazard, demand, and capacity respectively and U denotes uncertainty. Similarly, the dispersion parameters of the randomness in the demand and capacity are denoted by β_{DR}, and β_{CR}. Incorporating the uncertainty as defined above into Eq. 38, one can obtain the mean estimate of the limit state probability as follows:

$$E[P_t] = E[H(a^{\widetilde{C}})]\exp\left[\frac{1}{2}\frac{k^2}{b^2}\beta_{DR}^2 + \beta_{DU}^2 + \beta_{CR}^2 + \beta_{CU}^2)\right] \tag{40}$$

$$\text{in which } E[H(a^{\widetilde{C}})] = H(a^{\widetilde{C}})\exp\left[\frac{1}{2}\beta_{HU}\right] \tag{41}$$

In other words, effects of randomness and uncertainty are now combined. Note that the expected limit state probability is equal to the mean estimate of the hazard exceeding the median structural capacity multiplied by a correction factor that increases exponentially with the total uncertainty in the demand and capacity, and depends on the hazard and regression analysis parameters (k and b). The seismic hazard given in the USGS National Earthquake Hazard Maps is that of the mean hazard [Frankel *et al*, 1996] with regard to modeling uncertainty, Eq.40 is therefore compatible with the USGS hazard maps. Note also that the combination of uncertainty and randomness as shown in the second term of Eq. 40 is of the same form of the simple uncertainty analysis method of Eqs. 4 to 6.

To estimate the percentile values for a confidence interval estimate, in principal all uncertainty dispersions need to be considered when one extends the analysis to more than one uncertain parameter. In the FEMA/SAC procedure, it is assumed that the uncertainty in seismic hazard has been incorporated through the mean hazard curve in Eq.41. The confidence interval estimate is then obtained as function of the demand and capacity uncertainty using the mean hazard curve. The limit state probability corresponding to a percentile level q (probability of q not being exceeded) is given by

$$P_{q,t} = \widetilde{P}_t\exp[K_q\beta_L] \tag{42}$$

$$\text{in which, } \widetilde{P}_t = E[H(a^{\widetilde{C}})]\exp\left[\frac{1}{2}\frac{k^2}{b^2}(\beta_{DR}^2 + \beta_{CR}^2)\right] \tag{43}$$

$$\beta_L = \sqrt{\frac{k^2}{b^2}(\beta_{DU}^2 + \beta_{CU}^2)} \tag{44}$$

$$K_q = \Phi^{-1}(q) \tag{45}$$

$\widetilde{P}_t$ is the median (50%) value of P_t. Note that, the median estimate is the same as that considering only the randomness in the demand and capacity. The q percentile limit state probability is equal to $\widetilde{P}_t$multiplied by a factor depending on the uncertainty dispersion parameters and the percentile value. K_q is the standard normal variate value corresponding to this percentile value.

Continue with the example given in Eq.41 and now consider also the uncertainty. It is assumed that:

(1) $\beta_{HU} = 0.30$ in the hazard analysis (Eq. 31),
(2) $\beta_{DU} = 0.25$ describes the errors in the demand regression analysis as function of the spectral acceleration, and
(3) $\beta_{CU} = 0.28$ indicates the uncertainty in the estimate of capacity against collapse by the IDA procedure in Section 4.3.

The impact of different uncertainties on the incipient collapse probability can be illustrated as follows. The median incipient collapse probability of this steel building considering randomness only is 0.0208 (Eq.39). According to Eqs 40 and 41 when the seismic hazard analysis uncertainty is included in the form of mean hazard, this median value increases to

$$\widetilde{P}_{50} = H_{50}(0.86)\exp\left[\frac{1}{2}\times 0.3\right]e^{0.664} = 0.0241 \tag{46}$$

The mean estimate of the 50-year incipient collapse probability considering all uncertainties according to Eq. 40 is

$$\begin{aligned} E[P_{50}] &= H_{50}(0.86)\exp\left[\frac{1}{2}\times 0.3\right]\exp\left[\frac{1}{2}\frac{3^2}{1.2^2}(0.30^2+0.25^2+0.35^2+0.28^2)\right] \\ &= 0.0107\times 1.161\times 3.01 = 0.0374. \end{aligned} \tag{47}$$

It is seen that similar to the simple example in Section 2, the mean estimate is always higher than the median, reflecting the additional effect of uncertainties in demand and capacity. In this case, the probability almost doubles. Finally, when these uncertainties are accounted for in the form of a confidence level, then the probability is dependent on the confidence level. For example, the incipient collapse probability will not exceed the following value with 85% confidence (Eq. 42)

$$P_{85\%,\,50} = 0.0241\exp\left[1.03\times\sqrt{\frac{3^2}{1.2^2}(0.25^2+0.28^2)}\right] = 0.0616 \tag{48}$$

In other words, if in addition, a high confidence level is needed, the probability approximately triples that without considering the uncertainties. The above performance evaluation procedure has been applied to both pre- and post-Northridge steel moment frame buildings [Foutch 2000, Yun *et al* 2002, Lee and Foutch 2002]. For example, buildings designed in accordance with the 1997 NEHRP provisions and constructed with SAC pre-qualified connections have a confidence level of 90% of meeting the requirement of probability of incipient collapse being less than 0.02 in 50 years.

6 Reliability based design (FEMA/SAC procedure)

In view of the damages suffered in recent earthquakes, in the SAC/FEMA Joint Venture for Steel Buildings, a reliability-based and performance-oriented design has been developed where all randomness and uncertainty in the load and resistance are explicitly considered and accounted for [FEMA 350, 2000]. The critical issues related to such a statistical and reliability framework were reviewed in Wen and Foutch [1997]. The theoretical basis for the development of the design procedure can be found in Cornell *et al* [2002]. The final design format still retains the basic LRFD flavor with additional quantitative treatment of the effect of uncertainty. The results are being adopted in the AISC Seismic Provisions [Malley, 2002] and mostly likely will serve as a prototype for wider adoption in other codes and standards. The probability basis of this design procedure is briefly described in the following.

In the SAC/FEMA procedure, performance is checked at two levels, immediate occupancy (IO) and collapse prevention (CP) with associated target probability of 50% and 2% in 50 years respectively. Fig. 8 shows the performance checking of such a procedure. Assuming the probabilistic performance curves can be described by a distribution such as the lognormal, the two points on the curves allow one to describe and check the performance for a wide range of response. If the probability curve is higher than the target, stiffening, strengthening, or other mitigation measures are needed. Given the target probabilistic performance curve, the design is then to solve the inverse problem of finding the required structural capacity to meet the requirement. The simple closed form solution of the reliability analysis given above allows one to solve the inverse problem. Referring first to Eq.40 which gives the mean probability of a limit state over a time period, a reliability-based design is to determine the required structural median capacity, $\widetilde{C}$, to satisfy a prescribed target mean limit state probability, $E[P_t] = P_0$. This inverse problem can be solved as follows [Cornell *et al*, 2002]:

$$\left\{\exp\left[-\frac{1}{2}\frac{k}{b}(\beta_{CR}^2 + \beta_{CU}^2)\right]\right\}\widetilde{C} \geq \left\{\exp\left[\frac{1}{2}\frac{k}{b}(\beta_{DR}^2 + \beta_{DU}^2)\right]\right\}\widetilde{D}^{P_0} \tag{49}$$

It can be rewritten as

$$\phi\widetilde{C} \geq \gamma\widetilde{D}^{P_0} \tag{50}$$

ϕ is the capacity (resistance) factor and γ is the demand (load) factor. $\widetilde{D}^{P_0}$ = median demand corresponding to $S_a^{P_0}$, a spectral acceleration of exceedance probability of P_0. From Eq. 24, one obtains

$$\widetilde{D}^{P_0} = a(S_a^{P_0})^b \tag{51}$$

in which $S_a^{P_0}$ is solved from Eq.26. Note that from Eq. 51, smaller P_0 (higher reliability) gives larger $\widetilde{D}^{P_0}$ and from Eq. 49 larger randomness and uncertainty in the demand and capacity give larger γ and smaller ϕ leading to larger design capacity $\widetilde{C}$.

Continuing with the example problem in the previous section of the 3-story steel building, if the target 50-year incipient collapse (defined according to an IDA) probability is 2% what should be the median drift capacity against collapse? Note that since the target probability is lower than the current value of 0.0374 (Eq. 47), the structural

drift capacity against incipient collapse needs to be enhanced. From Eqs.49 and 50 the demand and capacity factors are calculated to be

$$\gamma = \exp\left[\frac{1}{2}\frac{3}{1.2}(0.30^2 + 0.25^2)\right] = 1.21 \tag{52}$$

$$\phi = \exp\left[-\frac{1}{2}\frac{3}{1.2}(0.35^2 + 0.28^2)\right] = 0.778 \tag{53}$$

and the median drift capacity corresponding to 2% in 50 years spectral acceleration exceedance probability is

$$\widetilde{D}^{2\%} = a(S_a^{2\%})^b = 0.06(0.698)^{1.2} = 0.039 \tag{54}$$

The required design median drift capacity against incipient collapse is therefore according to Eq.50

$$\widetilde{C} = \frac{1.21}{0.778} \times 0.039 = 0.0606 \tag{55}$$

Compared with the current median capacity of 0.05, an increase of 21% is required. The above design satisfies the target *mean* limit state probability considering all the randomness and uncertainty. Alternatively, one can also use a confidence level approach to set the design requirements as follows.

If a design criterion is that there must be a confidence level of at least q that the actual (but uncertain) limit state probability is less than the allowable value of P_0, then the formulation given in Eqs. 42 to 45 can be rearranged in terms of the factored capacity/demand ratio as follows

$$\lambda_{con} = \gamma\widetilde{D}^{P_0}/\phi\widetilde{C} \tag{56}$$

in which $\gamma, \phi, \widetilde{D}^{P_0}$, and $\widetilde{C}$ are defined in Eqs. 51to 53, λ_{con} is the demand/capacity ratio depending on the confidence level given by

$$\lambda_{con} = \exp\left[-K_x\beta_{UT} + \frac{1}{2}\frac{k}{b}\beta_{UT}^2\right] \tag{57}$$

$$K_x = \Phi^{-1}(q),\ q = \text{confidence level} \tag{58}$$

$$\beta_{UT} = \sqrt{\beta_{CU}^2 + \beta_{DU}^2},\ \text{total uncertainty in capacity and demand.} \tag{59}$$

For example, continuing with the example, if the target $P_0 = 2\%$ in 50 years and a confidence level of $q = 85$ % is desired. From Eqs. 57 to 59, one obtains

$$\begin{aligned} K_x &= \Phi^{-1}(0.85) = 1.04 \\ \beta_{UT} &= \sqrt{0.25^2 + 0.28^2} = 0.375 \\ \lambda_{con} &= \exp\left[-1.04 \times 0.375 + \frac{1}{2}\frac{3}{1.2}0.375^2\right] = 0.807 \end{aligned} \tag{60}$$

The required design median drift capacity is then determined from Eq. 56

$$\widetilde{C} = \frac{1.21 \times 0.039}{0.807 \times 0.778} = 0.075 \tag{61}$$

Compared with current design of 0.05, an increase of 50 % is needed to satisfy this design criterion.

7 Summary and conclusions

Recent developments of reliability analysis and reliability-based design of structures against earthquakes are given with emphasis on rigorous treatment of the large uncertainties in both demand and capacity including inherent variability and modeling errors. The analytical procedures are outlined; numerical examples on reliability analysis and reliability-based design are given to illustrate the concept and methodology. It is shown that both certainties play an important role in structural reliability and design decision. Modeling errors alone can double or triple the 50-yr incipient structural collapse probability and a 50% increase in structural capacity against incipient collapse. Application to current design procedures with specific consideration of the above two kinds of uncertainties are also mentioned.

References

Ang, A.H. and Tang, W.H. *Probability Concepts in Engineering Planning and Design-Vol.II Decision, Risks, and Reliability*, John-Wiley, 1984.

Boore, D.M. and Joyner, W.B. "Prediction of ground motion in North America." ATC-35-1, 6-1 to 6-41, 1994.

Collins, K.R., Wen, Y.K., and Foutch, D.A. "Dual-level design: a reliability-based methodology." *Earthquake Engineering and Structural Dynamics*, 25, pp.1433–1467, 1996.

Cornell C.A., Jalayer F., Hamburger, R.O., and Foutch, D.A. "Probabilistic basis for 2000 SAC Federal Emergency Management Agency steel moment frame guidelines." *Journal of Structural Engineering, ASCE*, 128(4), pp. 526–533, 2002.

Ellingwood, B.R., Galambos T.V., MacGregor J.G., and Cornell C.A. "Development of a probability based load criteria for American National Institute A58." NBS Special Publication 577, 1982a.

Ellingwood, B.R., Galambos T.V., MacGregor J.G., and Cornell C.A. "Probability-based load criteria: load factors and load combinations." *Journal of Structural Division*, ASCE, 108(5), pp. 978–997, 1982b.

Federal Emergency Management Agency (FEMA). *Recommended Seismic Design Criteria for New Steel Moment-Frame Buildings, FEMA 35, 2000.*

Frankel, A. "Simplified approach to incorporating uncertainty in the ground motion computation and mapping process." *ATC-35-2*, 8-1 to 8-38, 1997.

Frankel, A., Mueller, C., Barnhard, T., Perkins, D., Leyendecker, E.V., Dickman, N., Hanson, S., and Hopper, M. *National seismic-hazard maps: documentation*, *USGS Open-File Report* 96–532, 1996.

Liao, K.W., Wen, Y.K., and Foutch, D.A. "Evaluation of 3D Steel Moment Frames under Earthquake Excitations. I Modeling" *Journal of Structural Engineering*, ASCE, Vol 133, No. 3, pp. 462–471, 2007.

Liao, K.W., Wen, Y.K., and Foutch, D.A. "Evaluation of 3D Steel Moment Frames under Earthquake Excitations. II Reliability and Redundancy" *Journal of Structural Engineering*, ASCE, Vol 133, No. 3, pp. 471–480, 2007.

Luco, N. and Cornell, C.A. "Structure-specific scalar intensity measures for near source and ordinary earthquake ground motions." submitted to *Earthquake Spectr, 2002*.

Malley, J.O. "Update on US seismic design requirements for steel structures", *Proc. 7th US National Conference on Earthquake Engineering*, Boston, Mass. July 21–25, 2002.

Miranda E. and Bertero, V.V. "Evaluation of strength reduction factors for earthquake-resistant design." *Earthquake Spectra*, 10, pp. 357–379, 1994.

Nassar, A.A. and Krawinkler H. (1992). "Seismic design based on ductility and cumulative damage demands and capacities," *Nonlinear Seismic Analysis and Design of Reinforced Concrete Buildings*, ed. Fajfar, P. and Krawinkler, H., Elsevier Applied Sciences.

Park, Y.J. and Ang, A.H-S. "Mechanistic seismic damage for reinforced concrete." *Journal of Structural Engineering*, *ASCE*, 111(4), pp. 722–739, 1985.

Park, Y.J., and Ang, A.H-S., and Wen, Y.K. "Seismic damage analysis of reinforced concrete builidngs." *Journal of Structural Engineering*, ASCE, 111(4), pp.740–757, 1985.

SAC Steel Project. *Seismic Design Criteria for Steel Moment-Frame Structures, SAC CD01-, 2000*.

Senior Seismic Hazard Analysis Committee (SSHAC) *Probabilistic seismic hazard analysis: a consensus methodology, 1995*.

Shome, N, Cornell, A.C., Bazzurro, P., and Carballo, J.E. (1998). "Earthquakes, records, and nonlinear responses." *Earthquake Spectra*, 14(3), pp. 469–500, 1998.

Somerville, P.G., Smith, N., Punyamurthula, S., and Sun, J. "Development of ground motion time histories for Phase 2 of the FEMA/SAC Steel Project." *Report No. SAC/BD-97/04, SAC Joint Venture*, Sacramento, California, 1997.

Structural Engineering Association of California (SEAOC), *Vision 2000, Performance-based Seismic Engineering of Buildings, 1995*.

Vamvatsikos, D. and Cornell, C.A. "Incremental dynamic analysis." *Journal of Earthquake Engineering and Structural Dynamics*, 31(3), pp. 491–514, 2002.

Wang, C.-H. and Wen, Y.K. "Evaluation of pre-Northridge low-rise steel buildings-Part I, Modeling." *Journal of Structural Engineering*, *ASCE*, 126(10), pp. 1151–1160, 2000.

Wang, C.-H. and Wen, Y.K. "Evaluation of pre-Northridge low-rise steel buildings-Part II, Reliability." *Journal of Structural Engineering*, ASCE, 126(10), pp. 1161–1169, 2000.

Wen, Y.K. "Methods of random vibration for inelastic structures." *Applied Mechanics Reviews*, ASME, 42(2), pp. 39–52, 1989.

Wen, Y.K. *Structural Load Modeling and Combination for Performance and Safety Evaluation*, Elsevier, Development in Civil Engineering Series, Vol.31, 1990.

Wen, Y.K. and Foutch, D.A. "Proposed statistical and reliability framework for comparing and evaluating predictive models for evaluation and design and critical issues in developing such framework." *Report No. SAC/BD-97/03, SAC Joint Venture*, Sacramento, California, 1997.

Wen, Y.K. and Wu, C.L. "Uniform hazard ground motions for mid-America cities." *Earthquake Spectra*, 7(2), pp. 359–384, 2001.

Working Group on California Earthquake Probabilities. "Seismic hazards in southern California; probable earthquakes, 1994 to 2024." *Bulletin Seismological Society of America*, 85, pp. 379–439, 1995.

Chapter 14

Life-cycle cost and bridge management

Hitoshi Furuta

Kansai University, Takatsuki, Japan

ABSTRACT: The basic concept and advantage of Life-Cycle Cost (LCC) are discussed. The outline of bridge management system is illustrated and defects and problems to be overcome are clarified. Several suggestions are provided to realize a rational and economic maintenance program and to develop a practical bridge management system.

1 Introduction

Bridge maintenance is becoming more and more important all over the world. In Japan, a lot of highway bridges have been constructed over past 50 years, whose 40% bridges were constructed during the so-called highly developing period. In the coming ten years, the number of bridges more than 50 years old becomes four times of the present number, and further 10 years later it becomes 17 times. Most of these bridges are aging and suffering from damage, deterioration and environmental attack. The number of deteriorating bridges must increase in the near future.

This paper presents the current status of bridge maintenance and bridge management system. The outline of bridge management system is illustrated and defects and problems to be overcome are clarified. In Japan there are 150,000 bridges with spans of more than 15 meters are in use so that it requires an enormous budget to maintain them in satisfactory conditions. Furthermore, sufficient data, experienced maintenance engineers, special technologies, and good ordering system are needed.

In order to develop a bridge management system, the concept of Life-Cycle Cost (LCC) has been paid attention as a promising tool for the achievement of rational maintenance programs (Frangopol and Furuta, 2001). The basic concept and advantage of LCC are discussed. Several suggestions are provided to realize a rational and economic maintenance program and to develop a practical bridge management system.

2 Problems of bridge maintenance

2.1 *Present problems*

There are severe environments surrounding bridge maintenance as follows:

1) Reduction of investment for infrastructure
2) Increase of renewal investment
3) Increase of maintenance cost
4) Increase of stock of infrastructure

Considering the above situation, it is necessary to overcome the following problems:

1) Economic problems: Enormous budget is necessary to maintain a great number of infrastructures. For instance, there are 150,000 bridges (span length is greater than 15 m) in Japan.
2) Labor problems: It is difficult to collect a sufficient number of experienced maintenance engineers.
3) Lack of data: No satisfactory data is available due to old structures.
4) Technical problems: It is difficult to assess the integrity of existing bridges and to select appropriate repair methods, because each member has its own characteristic of deterioration.
5) System problems: Mismatching to the actual ordering system or estimating system.

2.2 *Solution of economic problems*

So far, initial construction cost has been considered as the most dominant economic factor. There has been no total insight that accounts for the maintenance cost.

Recently, Life-Cycle Cost (LCC) has been paid attention as a powerful and useful tool to achieve a rational maintenance program. This implies that a strategy with the minimum LCC can be an economic and safety plan from the view points of total economy. However, there still remain many problems to calculate the LCC, for example, asset evaluation, prediction of deterioration, effect of repair, etc.

3 Life-cycle cost (LCC)

3.1 *Definition of LCC*

In general, Life-Cycle Cost (LCC) is defined as

$$LCC = C_I + C_M + C_F \times P_f + C_R \tag{1}$$

where C_I: initial construction cost, C_M: maintenance cost, C_F: failure cost, P_f: failure probability, C_R: replacement cost. Note that C_M and C_R often include user cost. It is easily understood that P_f, C_M and C_F are difficult to estimate, because they involve various uncertainties. It is also difficult to estimate the maintenance cost due to the prediction of future deterioration.

3.2 *Application of GA to LCC optimization*

LCC optimization is to minimize the expected total cost which includes the initial cost involving design and construction, routine or preventive maintenance cost, inspection, repair and failure costs. The details are provided in (Frangopol et al. 1997). Cost of repair is calculated by using the rate of the initial bridge condition and the deteriorated condition, taking into account the effects of aging and corrosion, and all possibilities of repair based on an event tree. Failure cost is calculated based on the changing rate of reliability. Moreover, expected total cost is calculated by introducing the net discount rate of money.

LCC optimization is a nonlinear problem that includes integer and discrete variables. Therefore, it is necessary to apply a combinatorial optimization method to solve it. The purpose of LCC optimization is to find the most economical plan for inspection/repair. It is evident that a non-uniform interval of inspection/repair strategy is more economical than a uniform one (Frangopol et al. 1997). It is easily understood that the combination of inspection techniques with different detection capabilities in a strategy is more economic. Discrete variables are useful in determining when and how inspections and repairs have to be performed and what methods have to be used. Genetic Algorithm (GA) is a representative algorithm of combinatorial optimization methods (Goldberg, 1989). Using GA, it is possible to decide the number of lifetime inspections, the time of each inspection, and which inspection has to be used. Then, the time of repair is decided based on an event tree analysis.

The LCC optimization is reduced to the following mathematical programming problem:

$$\begin{array}{ll} \text{Minimize} & C_{ET} \\ \text{subject to} & P_{f,life} < P^*_{f,life} \end{array} \tag{2}$$

where C_{ET} is the expected total cost, $P_{f,life}$ is the lifetime probability of failure, and $P^*_{f,life}$ is the maximum acceptable lifetime probability of failure (Furuta et al., 1998).

4 Structural performance measures and multi-objective GA

4.1 *Structural performance*

LCC is one of the useful measures for evaluating the structural performance from another standpoint, which can reduce the overall cost and achieve an appropriate allocation of resources. In general, LCC optimization is to minimize the expected total cost which includes the initial cost involving design and construction, routine or preventive maintenance cost, inspection, repair and failure costs. Then, the optimal strategy obtained by LCC optimization can be different according to the prescribed level of structural performance and required service life. In this study an attempt is made to discuss the relationships among several performance measures and provide some good balances of them by using Multi-Objective Genetic Algorithm (MOGA).

4.2 *Formulation using Multi-Objective GA*

Multi-objective optimization problem has several objectives, and Pareto solutions are obtained as a set of solutions. GA evaluates the optimal solution by random and multiple-point searches. The Multi-Objective Genetic Algorithm (MOGA) is performed according to the following five steps (Goldberg, 1989):

Step 1: Generation of initial population.
Step 2: Crossover. Two-point crossover is used in this study.
Step 3: Mutation.
Step 4: Evaluation of fitness function.
Step 5: Selection.

Although several performance measures have been so far developed, the following four measures are considered here (Furuta et al. 2003a, Neves et al. 2004):

1) Safety Level: Safety level is defined in terms of structural capacity or durability.
2) Service Life: Service life means the prescribed expected period that the structure is used safely.
3) Life-Cycle Cost: Total cost including initial construction cost, maintenance cost, repairing cost, and replacement cost.
4) Condition Rate: Evaluation of damage state by inspection and defined in the range [0, 5]. 0, 1, 2, 3, 4, and 5 mean good, no damaged, slightly damaged, moderately damaged, severely damaged, collapse, respectively.
5) Reliability Index: Measure for structural safety obtained by reliability analysis.

The relationships among those performance measures are discussed by using MOGA. Paying attention to the two cases with three objective functions, the following two problems can be formulated:

Problem 1:
Objective function 1: Safety level should be maximized.
Objective function 2: Service life should be maximized.
Objective function 3: Life-cycle cost should be minimized.

Problem 2:
Objective function 1: Condition rate should be minimized.
Objective function 2: Safety index should be maximized.
Objective function 3: Life-cycle cost should be minimized.

By introducing MOGA, it is possible to obtain several available solutions that have different safety levels, service lives and LCC values, or condition rates, safety indices and LCC values (Furuta et al., 2003a, Furuta et al., 2004).

4.3 *Numerical examples*

A group of ten concrete highway bridges are considered in this study. The locations of all these bridges along the coast are indicated in Fig. 1. Maintenance management planning for ten consecutive piers and floor slabs (composite structure of steel girders and reinforced concrete (RC) slabs) is considered here. Each bridge has the same structure and is composed of six main structural components: upper part of pier, lower part of pier, shoe, girder, bearing section of floor slab, and central section of floor slab (Fig. 2) (Furuta et al., 2003b).

Environmental conditions can significantly affect the degree of deterioration of the structures and may vary from location to location according to geographical characteristics such as wind direction, amount of splash, etc. To take the environmental conditions into account, the deterioration type and year from completion of each bridge are summarized in Table 1.

In this study, environmental corrosion due to neutralization of concrete, chloride attack, frost damage, chemical corrosion, or alkali-aggregate reaction are considered as major deteriorations. The structural performance of each bridge component i is

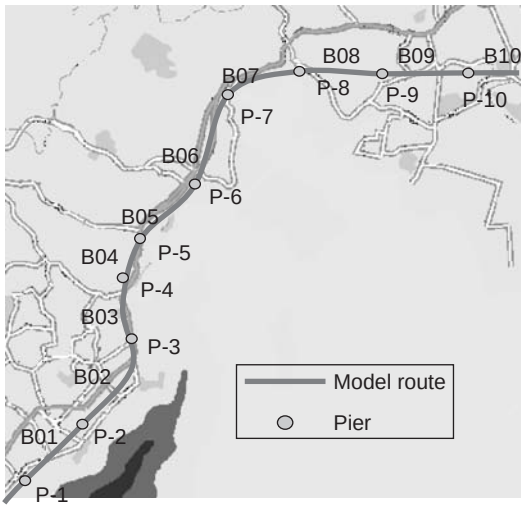


Figure 1 Locations of Ten Bridges in Japan.

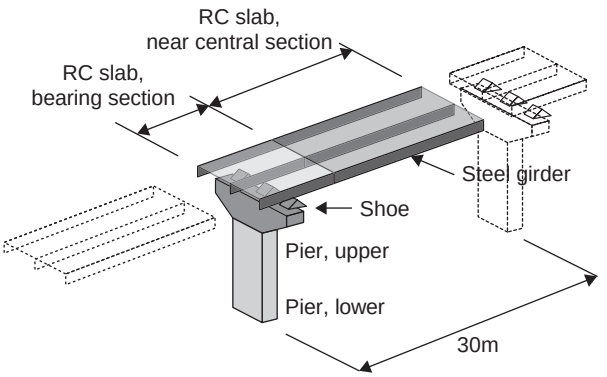


Figure 2 Main components of a bridge.

Table 1 Years from completion and type of deterioration caused by environmental conditions.

Bridge number	*Years from completion*	*Deterioration type*
B01	2	neutralization of concrete
B02	2	neutralization of concrete
B03	0	chloride attack (slight)
B04	0	chloride attack (medium)
B05	0	chloride attack (severe)
B06	0	chloride attack (medium)
B07	0	chloride attack (severe)
B08	1	chloride attack (medium)
B09	1	chloride attack (slight)
B10	1	chloride attack (slight)

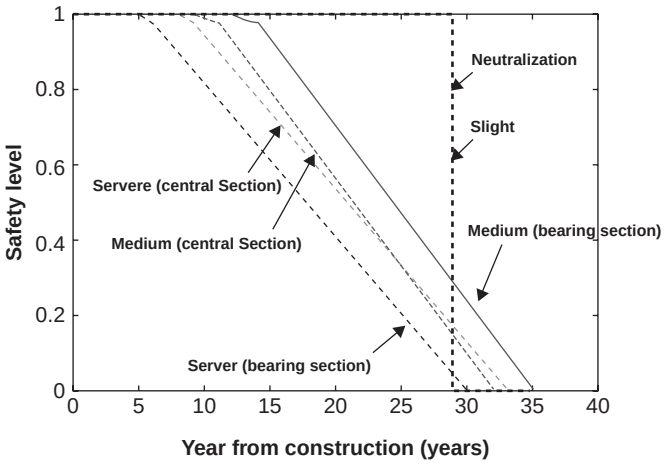


Figure 3 Typical performance of RC slabs.

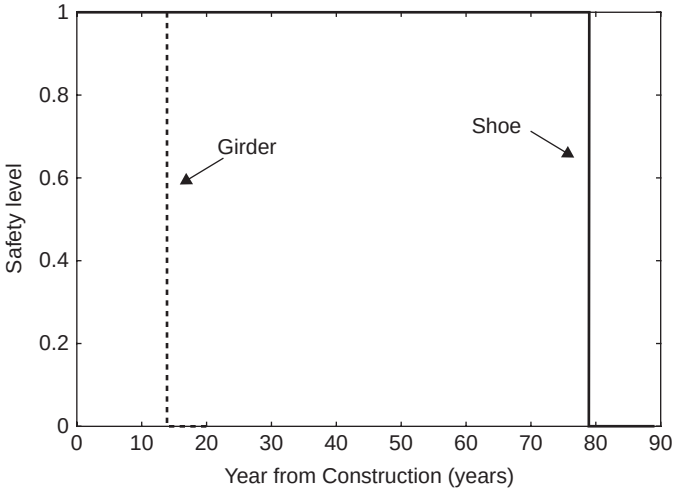


Figure 4 Typical performances of shoes and girders.

evaluated by the associated safety level (also called durability level) P_i which is defined as the ratio of current safety level to initial safety level.

Deterioration of a bridge due to corrosion depends on the concrete cover of its components and environmental conditions, among other factors. For each component, the major degradation mechanism and its rate of deterioration are assumed corresponding to associated environmental conditions. Figs. 3, 4, and 5 show the decreasing patterns of safety levels for RC slabs, shoes and girders, and piers, respectively. Average values are employed here as representative values for each level of chloride attack because the deteriorating rates can vary even in the same environment.

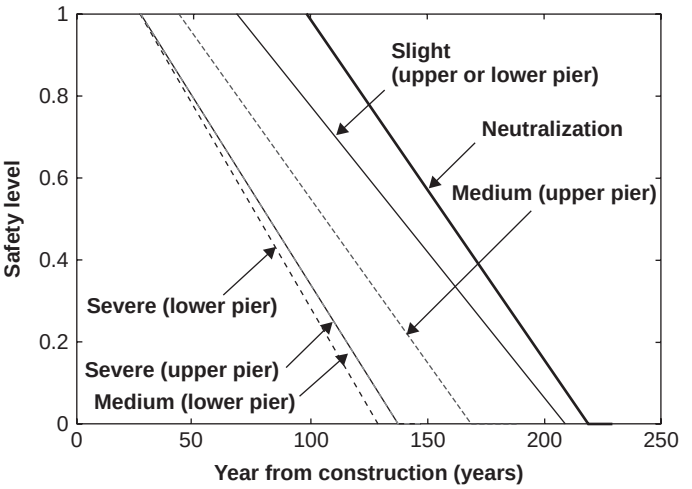


Figure 5 Typical performances of piers.

Table 2 Applicability and effects of repair methods.

Repair reinforcement	*Average effect*
Surface painting	Delays P_i decrease for 7 years
Surface covering	Delays P_i decrease for 10 years
Section restoring (Providing reinforcement)	Restores P_i to 1.0, and then allows it to deteriorate along the initial deterioration curve
Desalting (Re-alkalization)	Allows P_i to deteriorate along the initial deterioration curve
Cathodic protection	Delays P_i decrease for 40 years
Section restoring with surface covering	Restores P_i to 1.0, delays P_i decrease for 10 years, and then allows P_i to deteriorate along the initial deterioration curve

The decrease of RC slab performance is assumed to depend on corrosion. Hence, the safety level depends on the remaining cross-section of reinforcement bars. For shoe and girder, the major deterioration mechanism is considered to be fatigue due to repeated loadings. The decrease in performances occurs as the rubber bearing of shoe or paint coating of girder deteriorates. For pier, the major mechanism for deterioration is assumed to be only environmental corrosion. Thus the reduced performance of pier is expressed by the remaining section of reinforcement bars. The development of reinforcement corrosion is determined in accordance with Standard Specification for Design and Construction of Concrete in Japan (MLIT, 2002). In order to prevent deterioration in structural performance, several options such as repair, restoring, and reconstruction are considered. Their applicability and effects on each component are shown in Table 2.

In the implementation of MOGA, the GA parameters considered are as follows: number of individuals = 2000, crossover rate = 0.60, mutation rate = 0.05, and

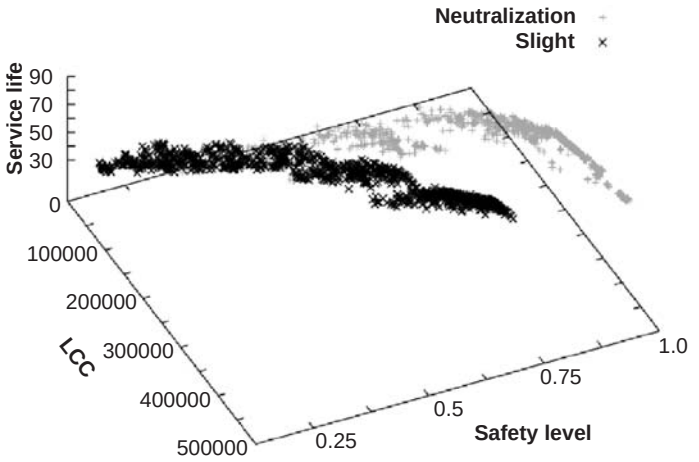


Figure 6 Results obtained by MOGA (1).

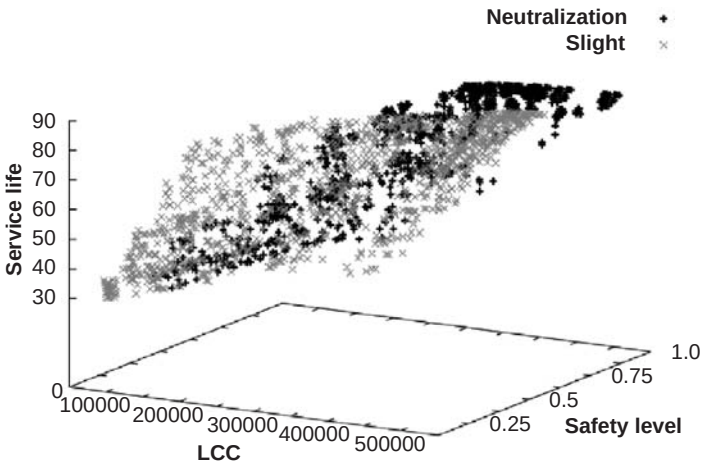


Figure 7 Results obtained by MOGA (2).

number of generations = 5000. Figs. 6 and 7 present the results obtained by MOGA. Each figure shows the comparison of the results of the neutralization environment and the severe chloride attack environment of a bridge from different points of view. From both figures, it is seen that the distributions of solutions under severe and neutralization environments are quite different. This is due to the fact that LCC is largely affected by the repairing methods employed. Figs. 8 and 9 show the relations between LCC and safety level. Fig. 8 presents the solutions obtained for the slight and neutralization environments, whereas Fig. 9 presents the results for medium and severe environments. From Fig. 8, it can be seen that the solution distributions are classified into three groups that have different cost-effectiveness points. These changing

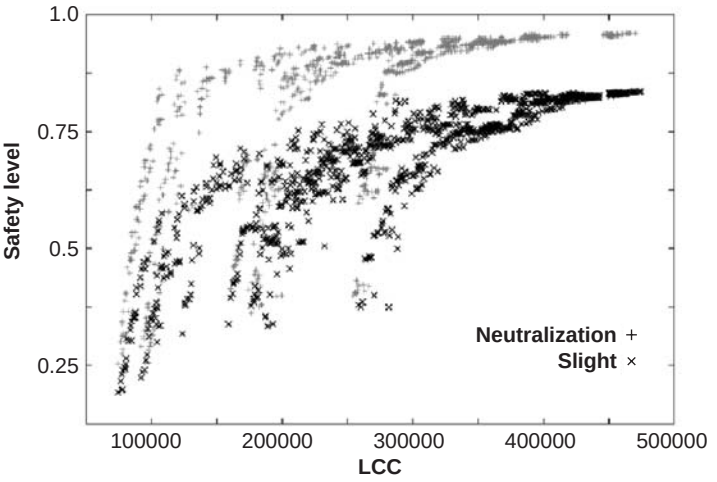


Figure 8 Relations between LCC and safety level for medium and severe environments.

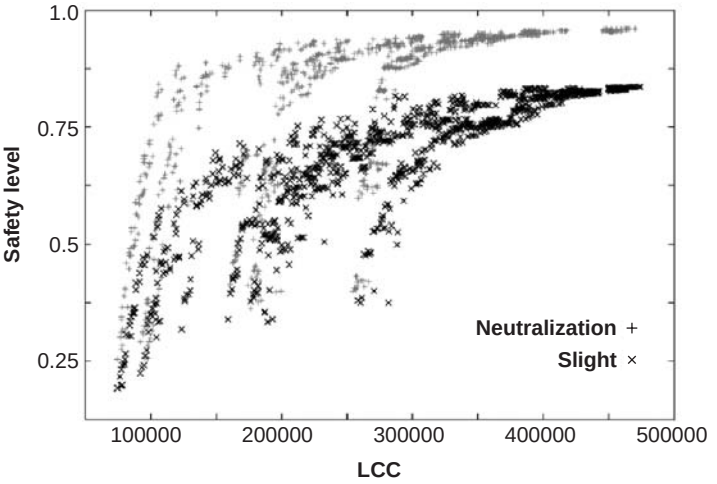


Figure 9 Relations between LCC and safety level for medium and severe environments.

points are 100 million yen, 200 million yen, and 280 million yen, respectively. The safety level can not be effectively improved even though the LCC is increased over the cost-effectiveness changing points. Comparing the solutions for the neutralization environment with the solutions for the slight chloride attack environment, it is obvious that these solutions show a different tendency so that it is necessary to be careful in establishing a maintenance plan. Different from Fig. 8, Fig. 9 indicates that the solutions have uneven envelopes and are concentrated into the center of their distribution. Figs. 10 and 11 present the relationships between LCC and service life, and Figs. 12 and 13 present the relationships between safety level and service life.

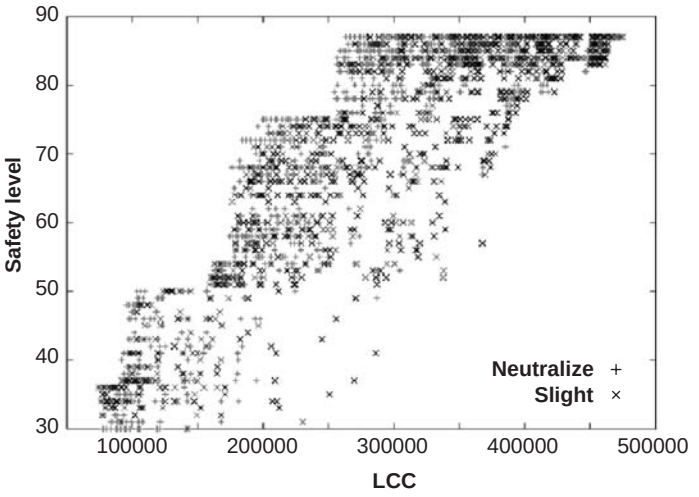


Figure 10 Relations between LCC and service life for neutralization and slight environments.

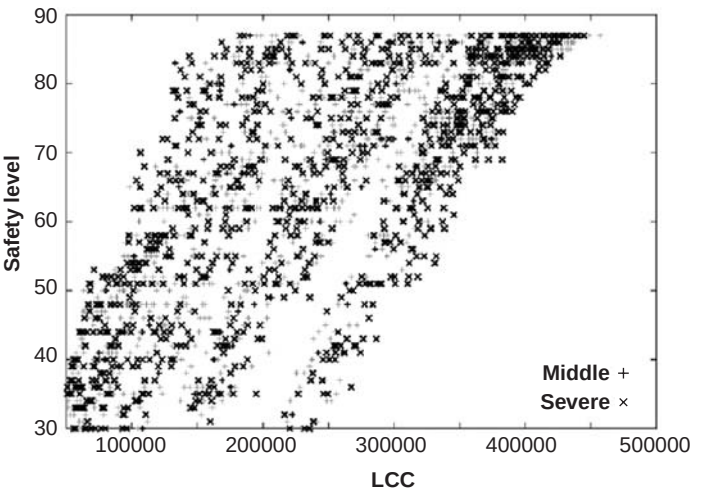


Figure 11 Relations between LCC and service life for medium and severe environments.

Figs. 14 to 16 present three representative maintenance plans. Figs. 14 and 15 show the rather short term (35 years) maintenance plans for the neutralization and slight chloride attack environments, respectively. The maintenance plan for the slight chloride attack environment requires 2.5 times LCC of that for the neutralization environment in order to maintain the same safety level. Accordingly, the repair methods become different, namely, the repair cost for RC slabs requires 2.3 times under the slight chloride attack environment. Comparing Fig. 15 and Fig. 16, it is obtained that although LCC for the severe environment becomes larger because of the severe damage, the difference

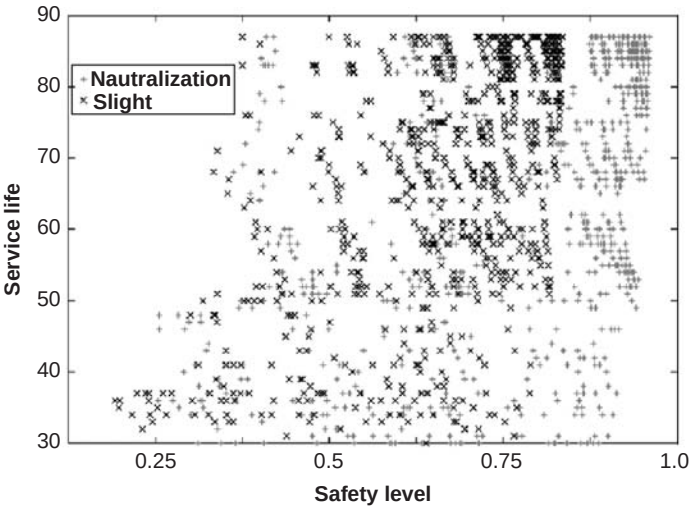


Figure 12 Relations between safety level and service life for neutralization and slight environments.

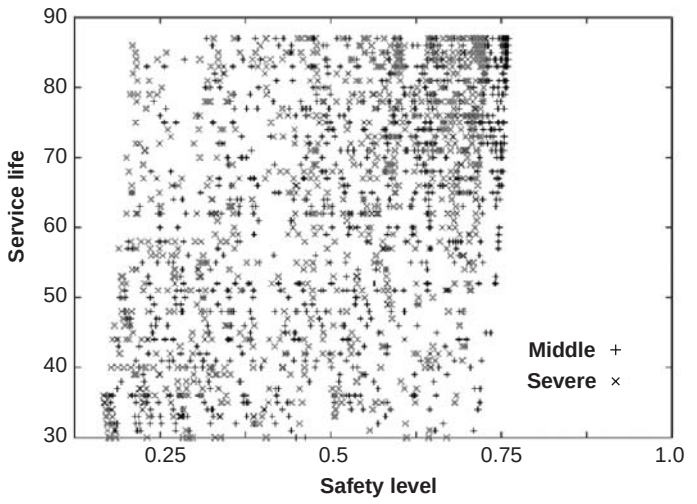


Figure 13 Relations between safety level and service life for medium and severe environments.

between them is not significantly large like the difference between neutralization and slight chloride attack environments.

5 New multi-objective GA to LCC optimization

Bridge maintenance planning has several constraints. In general, it is not easy to solve multi-objective optimization problems with constraints by applying a usual MOGA.

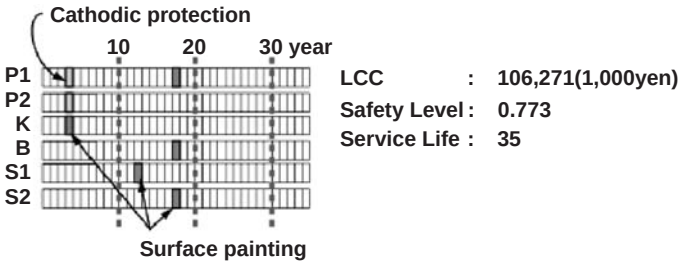


Figure 14 Short term maintenance plan for neutralization environment.

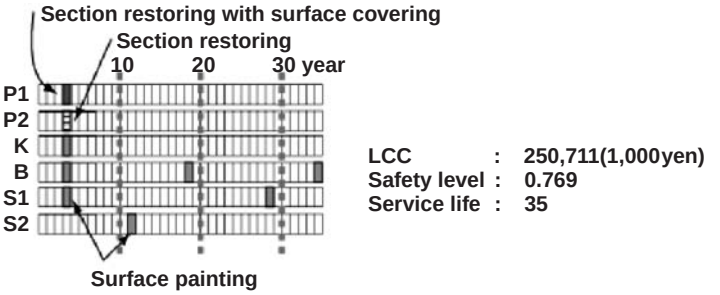


Figure 15 Short term maintenance plan for slight chloride attack environment.

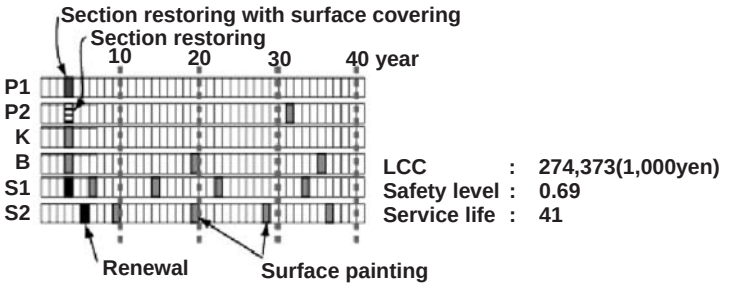


Figure 16 Short term maintenance plan for severe chloride attack environment.

5.1 *A new MOGA*

A new MOGA was developed by introducing the sorting technique into the selection process (Furuta et al. 2006). The selection is performed using so-called sorting rules which arrange the list of individuals in the order of higher evaluation values. Then, the fitness values are assigned to them by using the linear normalization technique. In general, if the fitness values are calculated directly according to the evaluation values, the differences among every individuals decrease so that the effective selection can not be done. In this study, the selection procedure is constructed coupling the linear normalization technique and the sorting technique. Using the evaluation values,

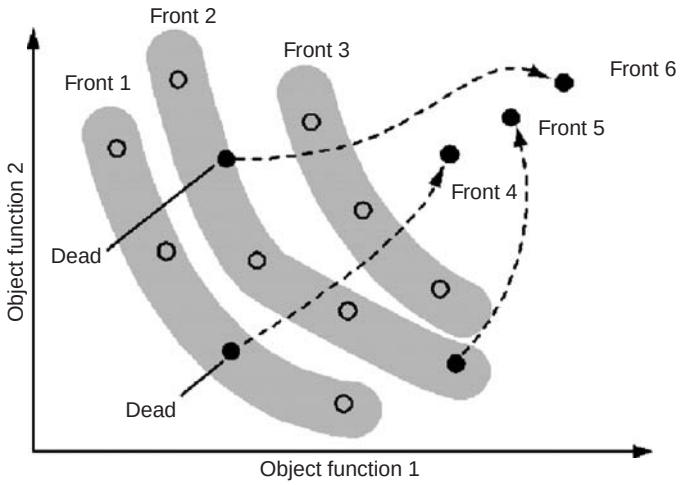


Figure 17 New sorting rules.

the individuals are reordered and given the new fitness values. Fig. 17 presents the process of the selection proposed here. The individuals of satisfying the constraints are arranged first according to the evaluation values and further the individuals of non-satisfying the constraints are arranged according to the degree of violating the constraints. Accordingly, all the individuals are given the fitness values using the linear normalization technique.

In order to apply the sorting rules to the multi-objective optimization problems, the non-dominated sorting method is used (Kitano, 1995). In the non-dominated sorting method, the Pareto solutions are defined as *Front1*. Then, *Front2* is determined by eliminating the *Front1* from the set of solution candidates. Repeating the process, the new *Front* is pursued until the solution candidates diminish. Further, the *Fronts* are stored in the pool of the next generation. If the pool is full, the individuals in the *Front* are divided into the solutions to survive or die based on the degree of congestion.

Then, the individuals are divided into the group of satisfying the constraints and the group without satisfying the constraints. The former is called as "alive individual", and the latter "dead individual". While the alive individuals are given the fitness values according to the evaluation values after the non-dominated sorting, the dead individuals are given the same fitness value. When implementing the non-dominated sorting, the Pareto *Front* may not exist at the initial generation, because a lot of dead individuals remain after the non-dominated sorting. Then, the dead individuals are arranged in the order of degree of violating the constraints and some of them are selected for the next generation. Thus, the multi-objective optimization problems with constraints are transformed into the minimization problem of violation of constraints. The elite preserve strategy is employed for the selection of survival individuals (Kitano, 1995).

When the generation progresses, alive individuals appear and then both the alive individuals forming the Pareto front and the dead individuals arranged in the order

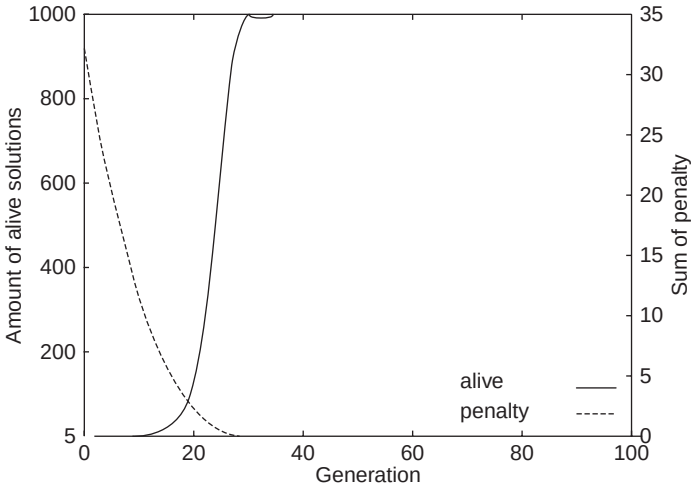


Figure 18 Evolution of new MOGA under neutralization environment (alive solutions and penalty).

of violation degree exist together. In this case, appropriate numbers of alive and dead individuals are selected for the next generation.

5.2 *Formulation using new MOGA*

In this study, LCC, safety level and service life are used as objective functions. LCC is minimized, safety level is maximized, and service life is maximized (Furuta et al., 2004). There are trade-off relations among the three objective functions. For example, LCC increases when service life is extended, and safety level and service life decrease due to the reduction of LCC. Then, multi-objective optimization can provide a set of Pareto solutions that can not improve an objective function without making other objective functions worse.

5.3 *Numerical example*

A group of ten highway bridges are considered again. Here, environmental corrosion due to neutralization of concrete, chloride attack, frost damage, chemical corrosion, or alkali-aggregate reaction are considered as major deteriorations. Four environmental conditions are taken into account; neutralization, mild, middle and severe environmental conditions.

In the implementation of the proposed new MOGA, the GA parameters considered are as follows: number of individuals = 1,000, crossover rate = 0.60, mutation rate = 0.05 and number of generations = 3,000.

Fig. 18 presents the results obtained by the new MOGA. This figure shows the evolution of the results from the 1st generation (iteration number) to the 100th generation. In Fig. 18, the sum of penalty started to decrease from the 1st generation to around

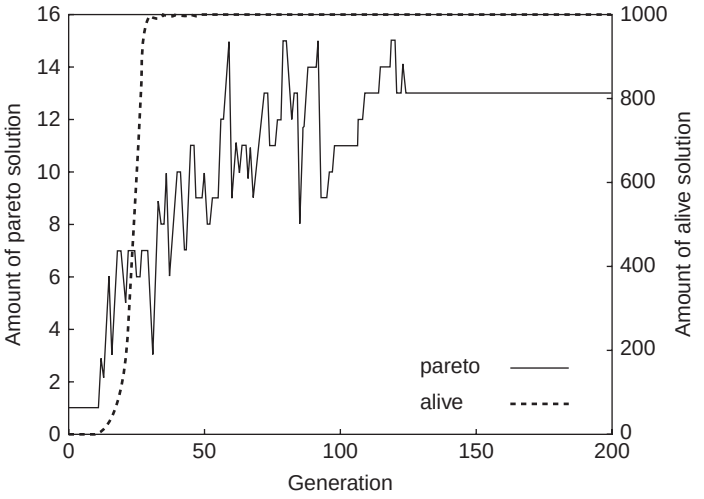


Figure 19 Evolution of new MOGA under neutralization environment (Pareto and alive solutions).

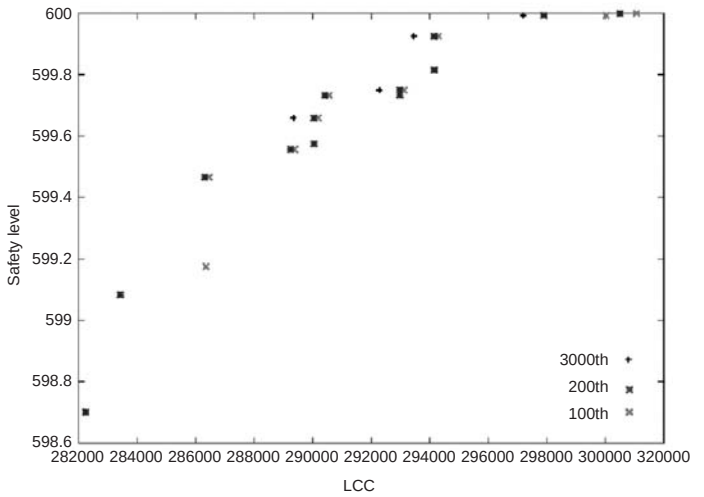


Figure 20 Evolution of new MOGA under neutralization environment (safety level and LCC).

the 30th generation. The 20th generation was starting to increase the amount of alive solutions.

In Fig. 19, the solutions at the 100th generation were not optimized. This means that the initial solutions can be generated uniformly. After the 120th generation, the solutions tend to converge to a surface, which finally forms the Pareto set as the envelope of all solutions. The number of solutions at the 3,000th generation is larger than that at the 100th generation. This indicates that the proposed new MOGA could

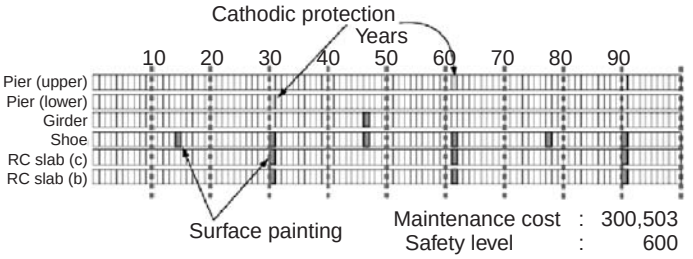


Figure 21 Maintenance scenarios for mild environment.

obtain various optimal solutions with different LCC values and safety levels. From Fig. 20, it is seen that the new MOGA can find out good solutions, all of which evolve for all the objective functions, and the final solutions are sparse and have discontinuity. In other words, the surfaces associated with the trade-off relations are not smooth. This implies that an appropriate long term maintenance plan cannot be created by the repetition of the short term plans.

Fig. 21 presents an optimal maintenance plan for the case of mild environment. It is noted that the simultaneous maintenance activities are considered to be preferable to reduce the LCC value. Consequently, it is confirmed that the proposed method can provide many useful solutions with different characteristics for determining an appropriate maintenance plan available for practical use. It is clear that LCC can be reduced by adopting simultaneous repair works. Finally, it is confirmed that the proposed method using linear normalization technique and sorting technique can provide many near-optimal maintenance plans with various reasonable LCC values, safety levels, and service lives. Note that it is quite difficult to obtain such near-optimal solutions by the current MOGA.

6 Life-cycle cost optimization considering seismic risk

6.1 *LCC with seismic risk*

In general, LCC is defined in terms of initial construction cost, maintenance cost, and replacement cost. However, in this study, LCC is defined as the sum of initial construction cost and seismic risk. As the initial construction cost, only piers are considered, because the sufficient data for the whole bridge is not available. Seismic risk includes both loss due to earthquake and user cost. Then, LCC considering seismic risk is calculated as

$$LCC = C_i + \sum P_d(a) \cdot C_d(a) \tag{3}$$

$$P_d(a) = P_b(a) \cdot P(DI, a) \tag{4}$$

where C_i: initial construction cost, $P_d(a)$: probability of seismic damage occurrence, $C_d(a)$: seismic loss, $P_b(a)$: earthquake occurrence probability, $P(DI,a)$: seismic damage probability, a: maximum acceleration, DI: damage index.

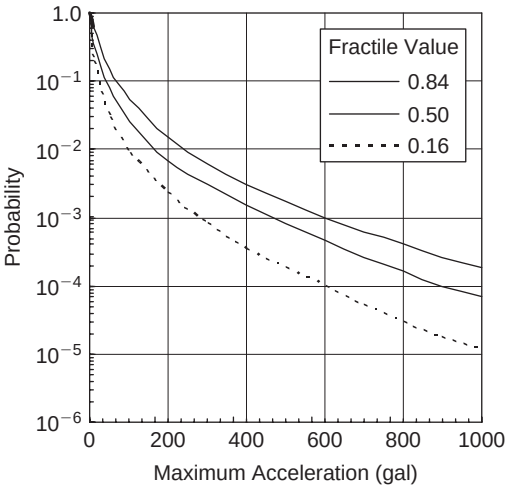


Figure 22 Earthquake hazard curve.

Table 3 Input earthquakes.

Earthquake type	*Name*	*M*	*Direction*
Type I	1968, Hyuga-Nada Earthquake	7.5	LG TR
	1994, Hokkaido-Toho Earthquake	8.1	TR
Type II	1995, Hyogo-Nanbu Earthquake	7.2	N-S E-W N27W

Eq. 4 provides the probability of damage occurrence due to the earthquake, which is the multiplication of earthquake occurrence probability with seismic damage probability. In this study, the earthquake occurrence probability is calculated by using the earthquake hazard curve shown in Fig. 22, and the damage probability is calculated by using the damage curve.

6.2 *Calculation of damage curve and definition of damage degree*

Seismic damage probability is defined in terms of the probability that a bridge pier shows each damage degree among the prescribed damage ranges. The damage curve is calculated here by using the dynamic analysis of the bridge pier.

Reinforced Concrete (RC) bridge pier is used to obtain the damage curve. The RC pier is designed according to Design Specification of Highway Bridges, Earthquake Version (MLIT 2002). Then, it is assumed that the ground condition is Type II and the importance of the bridge is B (MLIT 2002). Dynamic analysis is performed for bridges, in which a single mass and single degree of freedom model is used and Newmark β method is employed to do the dynamic analysis. It is assumed that the compression

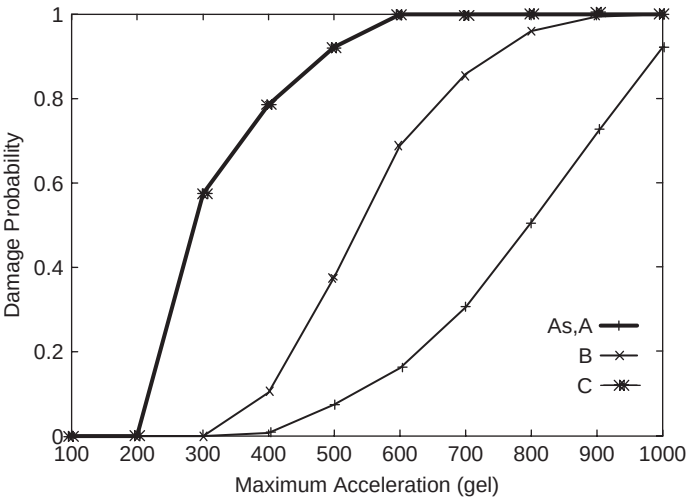


Figure 23 Damage probability and maximum accerelation.

strength of concrete is $f_c = 21(\mathrm{N/mm^2})$ and the reinforcing bars are SD295. As input earthquake wave, the ground condition of Type II is used and Type I and II earthquakes are used. Table 3 presents six input earthquakes. Using these conditions, the dynamic analysis is performed for 600 times for a RC pier. Several damage indices have been proposed so far (Park and Ang 1985, Shoji et al. 1997). However, in this research, damage degree is defined in terms of the maximum response displacement, horizontal force, and horizontal displacement of pier. The damage degree is categorized into five ranks such as As, A, B, C, and D that mean collapse, severe damage, partial buckling and deformation, deformation, and no damage or minor damage, respectively. In order to calculate the damage probability, it is necessary to determine the distribution function of damage degree (damage index) corresponding to the maximum earthquake acceleration. In this study, the log-normal distribution is assumed. When the distribution of damage degree is determined, the damage probability can be calculated as

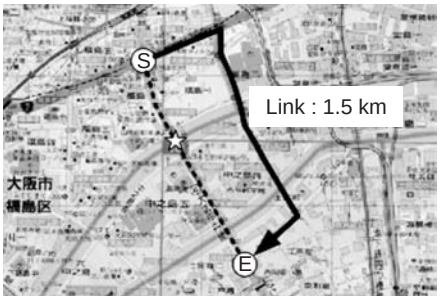
$$P(DI, a) = \int_a^b f_{DI}(x, a)dx \tag{5}$$

where [a, b] is the interval of each damage degree.

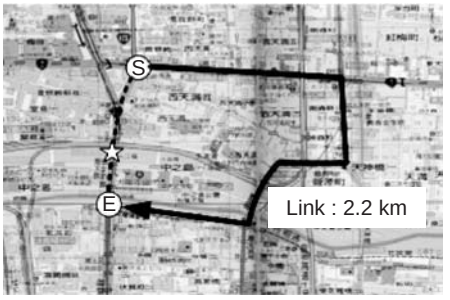
The damage probability is calculated for each damage degree and the results are plotted on a graph with the exceedance probability as the vertical axis and the maximum acceleration as the horizontal axis (JSCE 1996). Then, the damage curve can be obtained by combining them. Fig. 23 shows the computed damage curve.

7 LCC optimization for road network

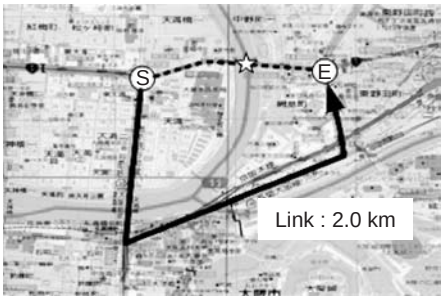
Practically, it is necessary to consider the effects of network to calculate LCC of bridge systems. It can be expected that although the effects of seismic risk are not large in the



Model 1. Tamae-bashi
Fukushima-ku, Osaka city



Model 2. Ohe-bashi
Kita-ku, Osaka city



Model 3. Sakuranomiya-bashi
Tyuo-ku, Osaka city

Figure 24 Three road network models.

case of a single bridge, they are large and important in the case of multiple bridges in a road network, because the user cost becomes quite large for the road network.

7.1 *Road network model*

For road networks, three network models (Model 1, Model 2, and Model 3) are employed, which are presented in Fig. 12. In these models, it is assumed that each network model includes a road passing over a river, and that traffics reach the destination through detours when some bridges can not be passed. Moreover, it is assumed that the traffic volume and the velocity have a relation shown in Fig. 24 (Kinki Branch, MLIT1999).

7.2 *User cost*

Here, user cost is defined in terms of the sum of the time cost and energy consumption cost due to the detour or closure of road. The cost associated with increasing driving time C_{UT} is calculated as the difference between (a) the cost associated with detour and road closure and (b) the usual cost (without detour and road closure).

$$UC_T = \alpha \cdot \{(Q \cdot T) - (Q_0 \cdot T_0)\} \tag{6}$$

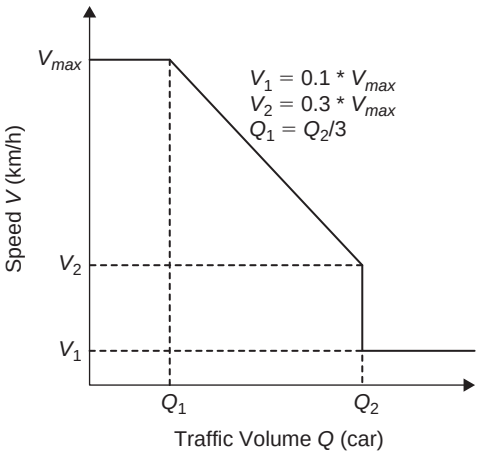


Figure 25 Relation between traffic volume and speed.

$$UC_C = \beta \cdot \{(Q \cdot L) - (Q_0 \cdot L_0)\} \tag{7}$$

where α: unit time cost, β: unit running cost, Q, T, L: detour traffic volume, running time, and link length at the time of road closure, Q_0, T_0, L_0: initial traffic volume, running time, and link length.

Using the data given in (Nihon Sogo Research Institute 1998) and assuming the ratio of small and medium trucks to be 10%, the unit time cost α is estimated as 82 Yen/car/min., and β is assumed as 18 Yen/car/km to 35 Yen/car/km. The restoring periods are assumed to be two months and two weeks for the damage (As, A) and B, respectively.

7.3 *Calculation of LCC*

Taking into account the discount rate, LCC is calculated as

$$LCC = C_i + \sum P_f(a) \cdot P(DI, a) \cdot \left\{ \frac{C_m(DI, a) + UC\,(DI, a)}{(1+i)^T} \right\} \tag{8}$$

where $C_m(DI,a)$: repair cost for each damage degree, $UC(DI,a)$: UC for each damage degree, i: discount rate, T: service life.

For each damage degree, restoring method and cost are presented in Table 4.

For the three road networks, LCC is calculated by assuming that the fractile value in the hazard curve is 0.5, discount rate is 0, service life is 100 years. Fig. 26 shows the calculated results, which indicate that there are important differences among the three networks, because of the differences in distances of detour and the initial traffic volumes. In the network with high traffics, seismic risk becomes 104,559,000 Yen that is 11 times the initial construction cost and 30 times the maintenance cost.

Table 4 Restoring method and cost.

Damage index	*Restoring method*	*Repair cost*	*Repair time*
As, A	Rebuild	120% of initial construction cost	2 month
B	Repair	73,000 Yen/1 m^2	1 month
C	Repair	35,000 Yen/1 m^2	2 weeks
D	No Repair		

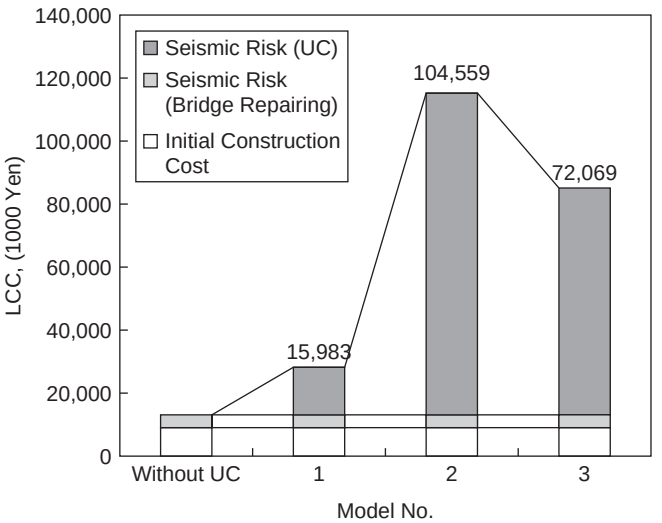


Figure 26 LCC of each model.

Comparing the case involving the user cost in the seismic risk and that without involving it, the seismic risk is only 1/3 of the initial cost when the user cost is not considered.

Paying attention to the damage probability curve in Fig. 23, it is evident that there is some difference in the damage probabilities according to the earthquake intensity. Fig. 27 shows the relation between the seismic risk and the maximum acceleration. This figure shows that it is obtained that the seismic risk decreases as the maximum acceleration increases. This is due to the facts that the bridge pier was designed to satisfy the requirement that the damage should be minor and the bridge function can be recovered in a short period. Therefore, the probabilities of damage B, C, and D become high.

The effects of damage degree on the seismic risk are examined. Fig. 28 presents the ratio of seismic risk corresponding to each damage degree, which implies that the damage degree C is 54% being the largest and the severe damages As and A have

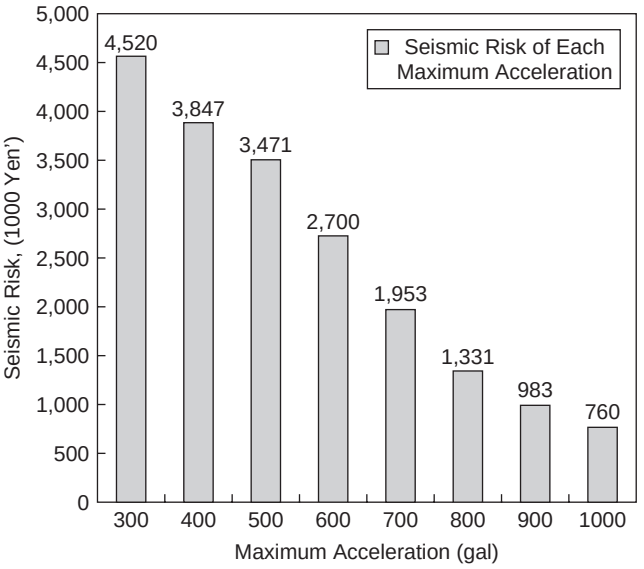


Figure 27 Seismic risk associated with each maximum acceleration.

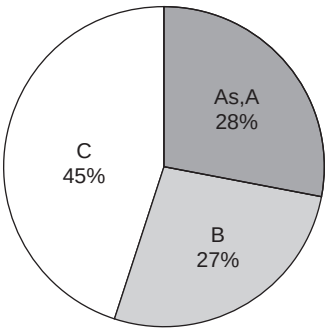


Figure 28 Relation between seismic risk and damage index.

small portions of 28%. This is due to the fact that while the occurrence probabilities of As and A become larger as the maximum acceleration increases, the earthquake occurrence probability decreases.

8 Bridge management system

Recently, the importance of structural maintenance has been widely recognized. In order to establish a rational maintenance program, it is necessary to develop a cost-effective decision-support system for the maintenance of existing infrastructures. As mentioned before, LCC is a useful concept in reducing the overall cost and achieving an appropriate allocation of resources.

In order to establish a bridge management system, it is inevitable to make the goal of the system clear. Namely, an appropriate maintenance level should be determined considering various factors. Bridge owner should account for the two standpoints of view; one is user's requirement and the other is the tax payer's requirement. The user requires safety and comfort. The safety includes structural safety and traffic safety, and the comfort includes drivability, noise and aesthetics. On the contrary, the tax payer requires the cost minimization for the maintenance. Consequently, the bridge maintenance level should be determined so as to maintain the bridges with the minimum LCC, guaranteeing their safety and comfort.

In order to calculate LCC, the following factors should be considered:

1) Inspection and investigation costs
2) Cleaning and maintenance costs
3) Emergency costs
4) Special costs, i.e., seismic retrofitting cost, etc.
5) Renewal costs

As shown previously, LCC is, in general, calculated based on the prediction of deterioration, in which the following conditions should be prescribed:

1) Deterioration curve
2) Present damage degree
3) Repair time (Maintenance level)
4) Repair method
5) Repair cost
6) Recovering rate after repair
7) Deterioration curve after repair

Although the above conditions need to be appropriately determined, it is quite difficult to achieve a measure for evaluating the comprehensive safety, integrity or damage state of bridges due to the complicated contribution of individual component to the overall function of bridge. Therefore, bridge authorities in Japan use the categorization of damage state such as "very severe", "severe", "moderate", "slight", and "very slight".

9 Present status and future of bridge maintenance in Japan

In Japan lots of bridges had been constructed during the last four decades. In the year of 2006, there are few bridges, i.e., 6%, being older than 50 years, however, the ratio can be estimated to rapidly increase such as 20% in the coming 10 years and 47% in 2026.

Ministry of Infrastructure, Land and Transport (MLIT), Japan decided to publish a bridge inspection manual in which inspection should be made every five years for about 20,000 bridges. To establish a strategic bridge maintenance program, Bridge Management System (BMS) has been being built. Local governments also have been attempting to develop their own BMS or Asset Management System (AMS).

AMS differs from BMS in the point that AMS focuses on the value of bridge as an asset. Namely, AMS aims to manage infrastructures, one of which is bridge, effectively

and efficiently from the long term vision. To establish a practical and useful BMS or AMS, it is necessary to overcome some problems like the variation of inspection results.

Furthermore, new technologies such as IT and sensing technologies can be available to make the bridge maintenance more effective. Enormous data can be handled by using the recent database technology, and graphical information can be utilized as well as text information. Advanced sensing and monitoring system can provide us with useful quantitative information. Nevertheless, the most important is the education for the inspectors or maintenance engineers to obtain satisfactory knowledge and experience. Sense of responsibility and passion of engineers may play an important role.

10 Conclusions

In this study an attempt was made to formulate LCC optimization based on GA and to discuss the relationships among several structural performance measures. In addition, seismic risk was introduced into the LCC optimization. The Multi-Objective Genetic Algorithm (MOGA) was adopted to successfully solve the large and complex combinatorial scheduling problems for the maintenance of damaged RC bridge structures.

By considering LCC, safety level, and service life as objective functions, it is possible to obtain the relationships among these three performance indicators and provide bridge maintenance management engineers with various maintenance plans and appropriate allocations of resources.

Based on the results presented in this study the following conclusions may be drawn:

1. Since the optimal maintenance problem is a very complex combinatorial problem, it is difficult to obtain reasonable solutions by the current optimization techniques.
2. Although Genetic Algorithm (GA) is applicable to solve multi-objective problems, it is difficult to apply it to large and very complex bridge network maintenance problems. By introducing the technique of Non-Dominated Sorting GA-2 (NSGA2), it is possible to obtain efficient near-optimal solutions for the maintenance planning of a group of bridge structures.
3. The Pareto solutions obtained by the proposed method show discontinuity. This means that the surfaces constructed by the trade-off relationships are not smooth so that an appropriate long term maintenance plan can not be created by the simple repetition of short term plans.
4. The optimal maintenance plans, especially repair methods, are quite different for the short term maintenance plan and the long term maintenance plan.
5. In the examples presented, the relation between safety level and LCC is non-linear. The increase of LCC hardly contributes to the improvement of safety level.
6. LCC can be reduced by adopting simultaneous repair works. The proposed new MOGA using linear normalization technique and sorting technique can provide many near-optimal maintenance plans with various reasonable LCC values and safety levels.
7. The damage degree is defined by using the maximum response displacement obtained by the dynamic analysis and the horizontal force and displacement of RC bridge pier.

8. Through the LCC calculation of several representative road networks, it is concluded that the difference of road network greatly affects the seismic risk.
9. Comparing the case with user cost and that without user cost, it is made clear that the effect of seismic risk is small unless user cost is considered.
10. Paying attention to the change of LCC according to the change of the maximum acceleration of earthquake, the seismic risk decreases, as the maximum acceleration increases.

References

Frangopol, D. M. & Furuta, H. (eds.), *Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, ASCE, Reston Virginia, 2001.

Frangopol, D. M., Lin, K.-Y., & Estes, A. C. , Life-cycle cost design of deteriorating structures, *Journal of Structural Engineering*, ASCE, Vol. 123 (10), pp.1390–1401, 1997.

Furuta, H., Frangopol, D. M. & Saito, M., Application of genetic algorithm to life-cycle cost maintenance of bridges, *Proc. of KKNN Symposium*, Seoul, Korea, 1998.

Furuta, H., Kameda, T., Fukuda, Y. & Frangopol, D. M., Life-cycle cost analysis for infrastructure systems: life-cycle cost vs. safety level vs. service life, *Proc. of 3rd Life-Cycle Cost Analysis and Design of Infrastructure Systems*, Lausanne, Switzerland, 2003a.

Furuta, H. & Koyama, K., Optimal maintenance planning of bridge structures considering earthquake effects, *Proc. of IFIP TC7 Conference*, Antipolis, France, 2003.

Furuta, H., Kameda, T., Nakahara, K., & Takahashi, Y., Genetic algorithm for optimal maintenance planning of bridge structures", *Proc. of GECCO,* Chicago, US, 2003b.

Furuta, H., Kameda, T. & Frangopol, D. M., Balance of structural performance measures, *Proc. of Structures Congress, Nashville, Tennessee, ASCE, May,* CD-ROM, 2004.

Furuta, H., Kameda, T. & Erami, M., A practical bridge management system using new multi-objective genetic algorithm, *Proc. of the 4th IABMAS Conference,* Porto, Portugal, 2005.

Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Company, Inc. 1989.

Japan Society of Civil Engineers, *Report on Damage by Hanshin Awaji Earthquake*, 1996. (in Japanese)

Kinki Branch, Ministry of Land, Infrastructure and Transportation, *Road Traffic Census,* 1999. (in Japanese)

Kitano, H. (ed.), *Genetic Algorithm 3*, Tokyo, Sangyo-tosho, 1995. (in Japanese)

Ministry of Land, Infrastructure and Transportation, *Design Specification of Highway Bridges*, Maruzen, 2002. (in Japanese)

Neves, L.C., Frangopol, D.M., & Hogg, V., Condition-reliability-cost interaction in bridge maintenance, *Proceedings of the Ninth International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP9*, San Francisco, California, July 6–9, Millpress, Rotterdam, 2, 1117–1122, 2003.

Nihon Sogo Research Institute, *Draft of Guideline for Evaluation of Investment on Road,* 1998. (in Japanese)

Park, Y. J., & Ang, A. H. S., Mechanistic seismic damage model for reinforced concrete, *Journal of Structural Engineering*, ASCE, Vol. 111, No. 4, 722–739, 1985.

Shoji, M., Fujino, Y. & Abe, M. Optimization of seismic damage allocation of viaduct systems, *Proc. of JSCE*, No. 563, 79–94, 1997. (in Japanese)

Chapter 15

Life-cycle cost analysis and design of civil infrastructures

Hyo-Nam Cho

Department of Civil & Environmental Engineering, Hanyang University, Ansan, Korea

ABSTRACT: Recently, the demand on the practical application of life-cycle civil engineering (LCCE) for tender design, optimum design and maintenance/management of civil infrastructures is rapidly growing unprecedently in civil engineering practice. Accordingly, in the 21st century, it is almost obvious that life-cycle civil engineering together with value engineering will become a new paradigm for all engineering decision problems in practice. However, in spite of impressive progress in the researches on the LCCE, the application of the LCCE in most countries is not yet mature enough to be pervasive into the practice for the design and management of major civil infrastructures. Thus, this paper is intended to review the current practice and suggest improved and practical approaches to the application of various LCC methodologies including performance controlled maintenance based on performance degradation, lifetime/reliability assessment, and subjective and updated uncertainty assessment for the LCC analysis and design of civil infrastructures. In addition, the LCC software systems developed for LCC analysis of highway and railroad bridge structures are reviewed with emphasis on integrated LCC system models that include the direct and indirect cost models, user friendly knowledge-based database incorporating statistical assessment and updating, reliability/risk assessment and optimization modules. Based on a few illustrative examples taken from real bridge structures, the practical application of LCC-analysis and LCC-effective optimum design of civil infrastructures are demonstrated in this paper. Also, it is shown that some advanced but practical methods such as performance degradation-based LCC analysis and equivalent LCC optimum design are applicable in practice for LCC analysis and design of civil infrastructures.

1 Introduction

Though the concept of Life-Cycle Cost (LCC) itself is not new, its effectiveness for planning, design, rehabilitation and maintenance/management of civil infrastructures is only recently becoming increasingly recognized, in practice, for the optimal decisions on all the engineering problems of planning, design, construction and asset management over the whole life span of a structure. For the decision problems as in the case of the LCC-effective maintenance of plant facilities, equipments, bridge decks, pavements, etc., the Life-Cycle Cost Analysis (LCCA) is relatively simple, and thus its practical implementation is rather straightforward. However, when it comes to major infrastructures such as bridge, tunnels, underground facilities, etc., the LCCA problem becomes extremely complex simply because time-variant degrading resistance and stochastic extreme load effects incur various failures related with strength, serviceability, durability, deterioration, and damage throughout the life span of a structure,

which, in turn, bring forth highly complicated cost and uncertainty assessment that often involves the lack of performance and cost data associated with various direct and indirect losses, and the absence of uncertainty data available for the assessment as well.

As a result, the LCC studies have been largely limited only to those relatively simple LCCA problems of planning or conceptual design for making decisions in optimal alternative design or rehabilitation. Accordingly, in the recent years, the researchers have pursued extensive studies on the LCC effectiveness mostly related to LCC models and frameworks for the optimal design criteria and optimal maintenance/management of civil infrastructures. Moreover, recently the demand on the practical application of LCC effective decisions in design and maintenance is rapidly growing unprecedently in civil engineering practice. And, thus the optimal decision criteria for design and maintenance of civil infrastructures are shifting from the initial cost-effective to the LCC effectiveness. Accordingly, in the 21st century, it is almost obvious that LCC together with Value Engineering (VE) will become a new paradigm for all decision problems in civil engineering practice along with the rapid development of information and intelligent computing technology.

The objective of this paper is to review conventional LCC analysis methods in practice and to suggest improved and practical LCC methodologies including performance controlled maintenance based on performance degradation, lifetime/reliability assessment and subjective and updated uncertainty assessment for the LCC analysis and design of civil infrastructures, mainly with the practical application to LCC-effective bridge design.

First, in this paper, a brief review of previous works is presented with the contemporary topics and issues from the aspects of practical implementation of LCC methodology. Next, the conventional procedure for LCC analysis in practice is briefly presented for the illustration of its basic concept and the comparison with advanced methods. And then, hierarchical definitions of LCC models are presented to categorize the level of LCC assessment applicable for the practical implementation of the LCC analysis and design of civil infrastructures. And then, an integrated LCC system model is introduced with an emphasis on data/uncertainty assessment and user-friendly knowledge-based database for its successful implementation. Finally, some illustrative examples are presented and discussed in order to demonstrate the practical application of the LCC analysis and design to various design problems.

In the last two decades, a number of researchers proposed methodologies for LCCA and Life-Cycle Cost assessment of design, maintenance and rehabilitation of civil infrastructures. Among various researches on LCCA, Veshosky's study (1997) is worth to mention, who performed the LCCA for the optimal decision on alternatives of bare deck with uncoated reinforcing bars, bare deck with epoxy-coated reinforcing bars, and protective Latex-Modified Concrete (LMC) overlay with uncoated reinforcing bars, etc., based on a survey of the opinions of bridge owners and engineers (Romano, 1997). Recently, Zimmerman (1997) established a guideline for evaluating maintenance costs and economic parameters including discount rates in LCCA. Noting that most of input data involves uncertainties in LCCA, Wall III and Smith (1998) proposed a probabilistic analysis approach for cost-effective pavement design emphasizing road user cost models.

As a result, in most of advanced countries various LCCA softwares for tender design or LCC-effective design of civil infrastructures have been developed to assess Life Cycle

Costing of various alternative construction materials, construction type, conceptual design, and maintenance of constructed facilities in practice. The most widely known software in public domain is BridgeLCC 1.0, which is a user-friendly LCCA program developed by the National Institute of Standards and Technology (NIST, 1996) to help engineers assess the cost effectiveness of new, alternative construction materials. The program uses a life-cycle costing methodology based on both of ASTM standard E-917 and a cost classification developed by NIST. However, this program has serious drawbacks in dealing with statistical uncertainties of input data. Therefore, recently NIST developed the new version, BridgeLCC 2.0 (http://www.bfrl.nist. gov/bridgelcc/), which makes up the weak points in BridgeLCC 1.0. However, this program may be possibly used only when individual users themselves could prepare most of input data. Therefore, it is necessary to construct a user-friendly knowledge-based DB which can be easily used in practice for LCCA.

More recently it is known that Hawk is developing a systematic LCCA model and software tool for cost-effective bridge design and maintenance (http://www4.nationalacademies.org/trb/crp.nsf/All+Projects/NCHRP +12-43), in which road user cost model and deterioration model are developed incorporating a Monte Carlo Simulation (MCS) technique. Also, recently in Korea, Cho, *et. al* (2002) have developed a standard LCCA software system for cost-effective bridge design, named PROLCC (PRObabilistic analysis-based Life-Cycle Cost analysis program), LCCSTEB(LCCA system for highway STEel Bridges), RAILBLCC(LCCA system for RAILroad Bridge). The essential features of these programs are such that they provide the benefits of utilizing statistical data available in Korea under the user-friendly operating system environment.

2 Basic procedure for LCC analysis

The basic procedure for LCC analysis itself is well known, standardized and widely used in practice for the LCC-effective decision making on design alternatives of various civil infra facilities. The basic procedure can be briefly illustrated as follows:

Step1. Establish Alternative Design Strategies.
The primary purpose of a LCCA is to quantify the long term implication of initial design decisions on the future cost of maintenance, repair and rehabilitation activities necessary to maintain some pre-established minimum acceptable level of service over some specified time period. The combination of initial design and necessary activities for maintenance, repair and rehabilitation represent a design strategy while the time horizon over which future cost are evaluated is referred to as the life cycle analysis period. One of the first steps in conducting a LCCA of alternative design is to identify the alternatives of design strategies for the LCC analysis period under consideration.

The LCC Analysis Period, in terms of the specified time period or the whole life span for LCCA should be sufficiently long enough to reflect long-term cost differences associated with reasonable design strategies. The LCC analysis period should generally always be longer than the design period except in the case of infrastructures with extremely long life span. The LCC analysis period should be long enough to incorporate at least minimum design life of alternatives in case of civil infrastructure like bridges.

Typically design strategies for each design alternative will have an expected initial design life, period to maintenance activities, and possibly a series of repair,

rehabilitation activities. The scope, timing, and cost of these activities must be identified. Depending on the initial design, the agency of infrastructure maintenance & management employs a variety of rehabilitation strategies to keep the infra-structures in functional condition. User costs & socio-economic losses due to reconstruction and rehabilitation strategies should also be identified.

Step2. Determine Performance Periods and Maintenance/Repair Activity Timing
Expected life performance of non-structural or structural elements, members, and subsystems for initial design and supporting rehabilitation activities has a major impact on LCCA results. It directly affects the frequency of agency intervention on the facilities, which in turn effects agency cost as well as user costs during period of maintenance and rehabilitation activities.

In most conventional approaches developed so far, maintenance scenarios are based on time-controlled maintenance (TCM) interventions for which the random application time is given as input data. On the contrary, performance-controlled interventions represent maintenance actions applied when a specific condition is reached. The application times of these interventions depend on both the target performance level and previous maintenance interventions. That is, the performance-controlled maintenance (PCM) scenario can reflect previous maintenance intervention, and thus, it seems that economical maintenance scenario which ensures a structure performance can be established (Kong & Frangopol, 2004).

LCC Analyst can determine specific maintenance scenario using historical records. It is obtained from facilities management or information system or internal records of agency.

When historical data is sufficient and available, the probability of each variable can be evaluated by using a simple frequency analysis. On the other hand, if the data is insufficient, engineering probability theories such as MCS or Bayesian approach could be used as uncertainty assessment methods. In other words, randomness and incompleteness of insufficient data is complemented by simulated or updated data. Also, if the data is not available at all, the occurrence probability of each variable may have to be assessed by subjective judgments based on experienced experts' opinion and knowledge. In this case, verbal descriptive (fuzzy) uncertainty assessment has been proven as a valuable tool for handling aleatory and epistemic uncertainties due to subjective estimates in decision making models (Cho et. al, 2002).

Step3. Estimate Life Cycle Costs.
Construction quantities and costs are directly related to the initial design and supporting maintenance strategy. As shown in Fig. 1., agency costs include all costs incurred directly by the agency over the whole life of the infrastructures. They typically include initial preliminary engineering, contract administration, construction supervision and construction cost, as well as future routine and preventive maintenance, replacement, retrofit and rehabilitation costs, and the associated administrative cost. User costs & socio-economic losses due to reconstruction and rehabilitation strategies are also estimated. More detailed information are not described herein but referred to Section 3.

Step4. Compute Net Present Value
Once all costs and their timing have been developed, future costs must be discounted to the base year and added to the initial cost to determine the net present value for the LCCA alternative. Accordingly LCCA software like BridgeLCC v2.0, LCCSTEB

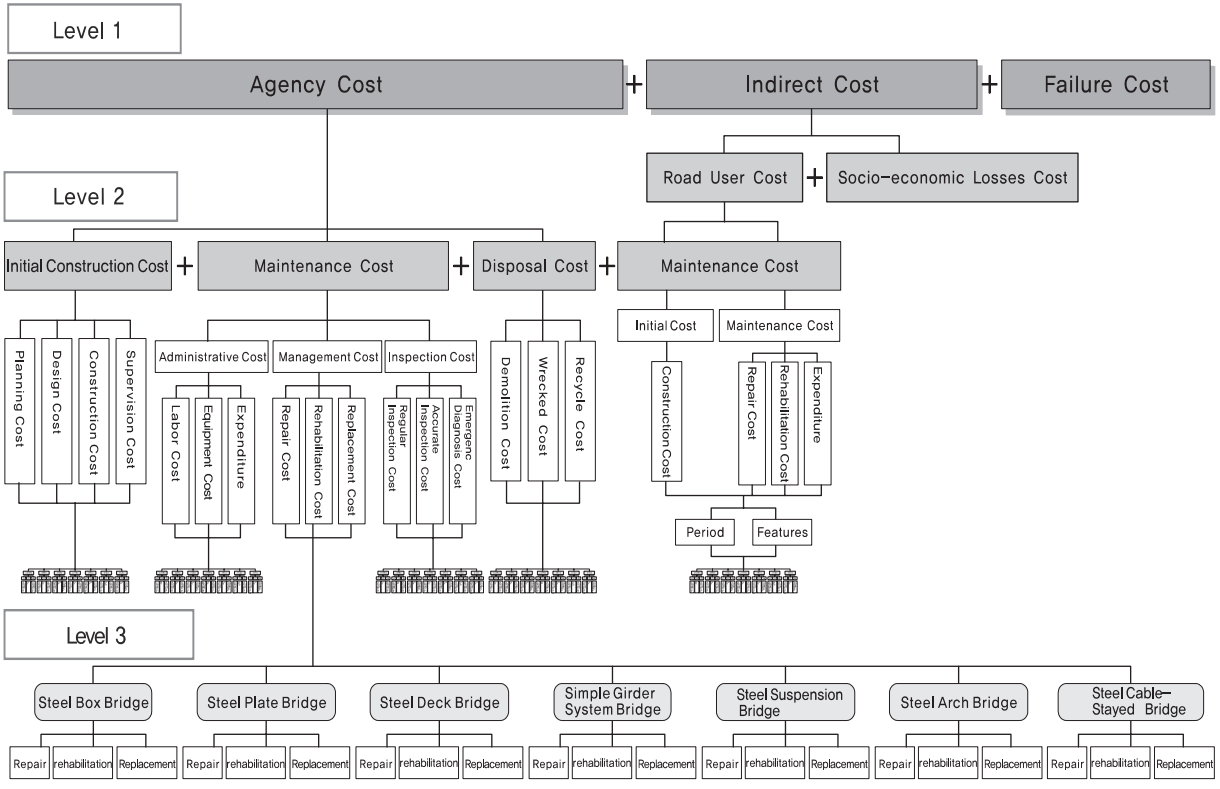


Figure 1 Cost classification (Lim, et al., 2005).

(developed in Korea), RAILBLCC(developed in Korea) use specified (or default) discount rate to estimate the net percent value of the LCC using data obtained from step 1, 2 and 3, and others general data.

Step5. Select the Best Alternatives.
Since PLCCA (Probabilistic LCCA) using MCS (Monte Carlo Simulation) is predominantly used in practice, once the distributions of LCC for each alternative are determined the best alternative can be selected for decision making based on the comparison of either some conservative percentile value of the LCC of each alternative (e.g. 90% confidence limit) or relative occurrence probabilities of each alternative LCC with respect to the minimum expected LCC.

3 LCC models

3.1 *LCC models for design-alternative analysis*

Theoretically, LCC analysis and design of civil infrastructures could involve very complex time-variant evaluation procedure mainly due to time variant performance degradation, loading history and operational environment. However, as it is shown in the previous section, practical approaches to LCCA are usually based on some simplified form of LCC model. In general, this simple LCCA can be divided into DLCCA (Deterministic LCC Analysis) and PLCCA (Probabilistic LCC Analysis). Both approaches are applicable to LCC analysis of infrastructures for tender design and/or as an essential part of design VE works. However, these days, DLCCA is not any more used in practice simply because this model only uses a simple deterministic estimate for each cost item without considering uncertainties of input variables in a LCCA. In other words, only one deterministic result is evaluated by using only one likelihood expected or representative value for each of time and cost parameters of the model in conducting DLCCA. And thus, often, stake-holders seize upon the uncertainties associated with LCCA inputs and vigorously debate the validity of the results or the decision on the alternatives.

Recently, as it is shown in Section 2, PLCCA is predominantly used in the tender design for bids, as the method that utilizes MCS in the treatment of uncertain input variables (type of distribution, most likelihood expected value, and coefficient of variation, etc.) to generate probabilistic results. Therefore, it is recommended in FHWA (Walls III and Smith, 1998) that this PLCCA approach is more reasonable and scientific than DLCCA, which is usually analyzed by Monte Carlo simulation of the whole range of possible outcomes using occurrence probability of economic input factors. Similar to the inputs, the results of PLCCA are visually presented in the form of a probability distribution. Thus, more information such as distribution of LCC, and the chance (probability) for becoming optimal alternative, etc. can be acquired.

Therefore, a general form of LCC model, corresponding to the cost item as shown in Fig. 1., may be formulated as follow:

$$E[C_T^{PV}(\vec{X})] = C_I^{PV}(\vec{X}) + E[C_M^{PV}(\vec{X})] + E[C_D^{PV}(\vec{X})] \tag{1-a}$$

$$E[C_D^{PV}(\vec{X})] = \sum_{k=1}^{K} E[C_{DSP_k}^{PV}(\vec{X})] + E[C_{REC_k}^{PV}(\vec{X})] \tag{1-b}$$

where $E[C_T^{PV}(\vec{X})]$ = total expected LCC which are functions of the vector of design parameters $\vec{X}$ in present worth; $C_I^{PV}(\vec{X})$ = initial investment cost in present worth; $E[C_M^{PV}(\vec{X})]$ = expected lifetime maintenance cost in present worth; $E[C_D^{PV}(\vec{X})]$ = disposal costs that consist of the sum of demolition costs $E[C_{DSP}(\vec{X})]$ and recycling costs $E[C_{REC}(\vec{X})]$, which depend upon design parameters $\vec{X}$ associated with component k quantities of infrastructures.

3.1.1 *Initial cost*

The initial investment costs involves in design and construction of the structural and non-structural components of the structures such as planning and design, construction of foundation, superstructure, substructure and accessories in the case of bridges. Thus, in general, the initial cost may be expressed in the following form:

$$C_I^{PV}(\vec{X}) = (C_{Des}(\vec{X}) + C_{Con}(\vec{X}) + C_{Tes}(\vec{X})) \cdot i(t) \tag{2}$$

where $C_{Des}(\vec{X})$ = planning and design cost; $C_{Con}(\vec{X})$ = construction cost; $C_{Tes}(\vec{X})$ = testing cost; and $i(t)$ = discount rate function, which can be expressed as $1/(1+q)^t$; q = discount rate; and t = time at which initial investment costs is implemented.

The construction cost should include all the costs of labor, materials, equipments, construction site management and quality control involved in the actual construction of civil infra-facilities. The construction costs can be expressed in terms of the initial cost, which may be approximately estimated at about 90% of the initial costs in the case of bridge, and the design cost and proof-load testing cost are also typically assumed in terms of a percentage of construction cost.

3.1.2 *Expected maintenance cost*

Since the maintenance costs over a life span may be high, depending upon the quality, type, location, and environment of design and construction of facilities, the costs must be carefully regarded as one of the important costs in the evaluation of LCC.

In general, the maintenance costs can be classified in terms of the costs which are related to design parameters and maintenance strategy. Thus, the expected maintenance costs may be formulated as follows:

$$E[C_M^{PV}(\vec{X})] = E[C_M^{PV}(\vec{X})] + E[C_{INS}^{PV}(\vec{X})] + E[C_{RM}^{PV}(\vec{X})] \tag{3-a}$$

$$E[C_{RM}^{PV}(\vec{X})] = \int_0^L [C_{DRM}(\vec{X}) + C_{IDRM}(\vec{X})] \cdot P_{RM}(t) \cdot i(t)dt \tag{3-b}$$

where $E[C_{PM}^{PV}(\vec{X})]$ = expected ordinary repair maintenance cost; $E[C_{INS}^{PV}(\vec{X})]$ = inspection/diagnosis costs which are function, expressed in terms of the facility types, scale, and construction year, etc.; $E[C_{RM}^{PV}(\vec{X})]$ = expected repair cost; $C_{DRM}(\vec{X})$, $C_{IDRM}(\vec{X})$ = direct and indirect repair cost, respectively; $P_{RM}(t)$ = repair probability

of a facility at time t, which may be estimated based on performance degradation uncertainties; and L = life span.

3.1.3 *Expected rehabilitation cost*

Expected rehabilitation costs may arise as a result of annual probability of failure or damage of various critical limit states that may occur any time during the life span of an infrastructure. Even though the rehabilitation costs do not have to be considered under normal circumstances, these costs should still be considered in an economic analysis as insurance costs (Cho, 2002). The expected rehabilitation costs can be obtained from direct/indirect rehabilitation costs and annual probability of failure for failure limit state i considered in the design as follows:

$$E[C_F^{PV}(\vec{X})] = \int_0^T \sum_{i=1}^{L} [(C_{DR}(\vec{X}) + C_{IR}(\vec{X})) \cdot P_{F_i}(\vec{X}, t|T)] \cdot i(t)dt \tag{4-a}$$

$$C_{IR}(\vec{X}) = C_U \cdot t_i(\vec{X}) + C_H \cdot r_i(\vec{X}) + C_E \cdot t_i(\vec{X}) \tag{4-b}$$

where i = index for a failure limit state; $C_{DR}(\vec{X})$ = direct rehabilitation cost for failure limit state; $P_{F_i}(\vec{X}, t|T)$ = updated annual probability of failure at any time t (i.e., annual probability that the failure will occur during time interval t conditional on updated loads or resistance); $C_{IR}(\vec{X})$ = indirect rehabilitation cost for failure limit state; C_U, C_H, C_E = loss of user cost, contents or fatality and injury losses, and indirect socioeconomic losses, respectively; $r_i(\vec{X})$ = increased accident rate during rehabilitation activities; and $t_i(\vec{X})$ = period of rehabilitation activities.

The direct rehabilitation cost $C_{DR}(\vec{X})$ and the loss of contents or fatality and injury loss cost C_H could be evaluated in a relatively easy way. In case of Korea, the direct rehabilitation costs are estimated based on the various sources available, such as Construction Software Research's price information (CSR, http://www.csr.co.kr), opinions of the experts, and also obtained from various references (OCFM, 2002; KISTEC, 2000). And also the loss of contents or fatality and injury losses cost C_H are evaluated based on the research results of Korea Transport Institute, using the human capital approach of the traffic accident cost data (KOTI, http://traffic.metro.seoul.kr).

3.1.4 *Indirect cost*

For an individual structure like a building structure, it can be argued that only the owner's cost may be relevant and thus it might have a minor influence on public user cost or socio-economic losses. However, when it comes to infrastructure such as bridge, tunnel, water delivery, and underground facilities, etc., the situation becomes completely different precisely because those infrastructures are primary public investments that provide vital service to entire urban areas. Thus, the indirect costs accruing to the public user of these infrastructural systems should also be accounted for. Herein, only the models for highway and rail road bridges as for specific application examples are presented in the following.

3.2 *Indirect cost models*

As aforementioned, each LCC model involves the assessment of direct and indirect costs. In particular, indirect costs deserve further specific study on each different kind of infrastrucrures. For rational assessment of indirect costs of bridge structures, indirect cost models of user cost and socio-economic losses applicable for highway and railroad bridges are suggested in this paper. Indirect cost of highway bridges may be expressed in terms of the road user costs, and socio-economic losses.

3.2.1 *User cost model*

In general, road user costs consist of 5 major cost items namely, vehicle operating costs, time delay costs, safety and accident costs, comfort and convenience costs, and environmental costs. Among the items, time delay costs and vehicle operating costs have been generally considered as major cost items of the road user cost (Lim, 1999; Cho et. al., 2001). To evaluate the rational road user costs, the essential factors such as traffic network, the location of bridge, and the information on repair (e.g., work zone condition, detour rate, the change of traffic capacity of traffic network, etc.) must be considered. Then, for example, the road user cost for highway bridge may be expressed as follows (Cho et. al, 2004):

$$C_U^{highway} = C_{TDC}^{highway} + C_{VOC}^{highway} \tag{5-a}$$

$$C_{TDC}^{highway} = \left[\begin{array}{l} \left\{ \sum_{j=1}^{J} n_{P_{0j}} \cdot T_{0j} \cdot u_{1_{0j}} \right\} \cdot \left(1 - \sum_{i=1}^{I} r_i \right) \cdot \Delta t_{d0} + \\ \sum_{i=1}^{I} \left\{ \sum_{j=1}^{J} r_i \cdot n_{P_{0j}} \cdot T_{0j} \cdot u_{1_{0j}} + n_{P_{ij}} \cdot T_{ij} \cdot u_{1_{ij}} \right\} \cdot \Delta t_{di} \end{array} \right] \tag{5-b}$$

$$C_{VOC}^{highway} = \left[\begin{array}{l} \left\{ \sum_{j=1}^{J} T_{0j} (u_{2j} + u_{4_{0j}} l_{d0}) \right\} \cdot \left(1 - \sum_{i=1}^{I} r_i \right) \cdot \Delta t_{d_0} + \\ \sum_{i=1}^{I} \left\{ r_i \cdot \sum_{j=1}^{J} (T_{0j} \cdot u_{2j}) + T_{ij} \cdot u_{2j} \right\} \cdot \Delta t_{d_i} + \\ \sum_{i=1}^{I} \left\{ r_i \cdot \sum_{j=1}^{J} \left[T_{0j} \cdot (u_{3_{ij}} l_{d_i} - u_{4_{ij}} l_{d0}) + T_{ij} \cdot u_{3ij} \right] \right\} \cdot \Delta t_{d_i} \end{array} \right] \tag{5-c}$$

$$\Delta t_{d_i} = \frac{l_{d_i}}{v_{d_{wi}}} - \frac{l_{d_i}}{v_{d_{ni}}}, \quad \Delta t_{d_0} = \frac{l_o}{v_{o_w}} - \frac{l_0}{v_{d_n}} \tag{5-d}$$

where $C_U^{highway}(\vec{X})$ = user cost of highway bridges; $C_{TDC}^{highway}(\vec{X})$, $C_{VOC}^{highway}(\vec{X})$ = time delay and vehicle operating costs; o = an index for original route in network i = an index for route in network; j = an index of types of vehicles which should be classified into those for business or non-business such as owner car for business, owner car for non-business, taxi, bus for business, bus for non-business, small truck, and large truck etc.; $n_{o_{ij}}, n_{P_{ij}}$ = number of passengers in vehicle; T_{oj}, T_{ij} = Average Daily Traffic Volume

(ADTV); l_o, l_{d_j} = the route length of bridge route (the route including bridge) and detour route; r_i = detour rate form original route to i-th route; $u_{1_{ij}}$ = average unit value of time per the user; u_{2j} = average fixed operator wages for each type of vehicle j; $u_{3_{ij}}$ = average unit fuel cost per unit length on the detoured route; $u_{4_{ij}}$ = the average unit fuel cost per unit length on the original route; $\Delta t_{d_0}, \Delta t_{d_i}$ = the additional time delay on the original route and i-th route; v_{o_n}, v_{o_w} = the average traffic speed on the original route during normal condition and repair/rehabilitation activity; and $v_{d_{ni}}, v_{d_{wi}}$ = the average traffic speed on detour route during normal condition and rehabilitation activity, respectively.

However, the user cost of railroad bridges is completely different from that of highway bridges, because it may be easily realized that no detour lines usually exist for a railway bridge line, and train (vehicle) operating costs should not be included in the user cost since trains are owned not by users but by public corporations. So the user cost of railroad bridges could be estimated only with time delay costs. Then, the user cost of railroad bridges may be expressed as follows:

$$C_U^{railroad} = C_{TDC}^{railroad} = \sum_i \sum_j n_{P_{ij}}^{railroad} \cdot T_{ij}^{railroad} \cdot u_{ij}^{railroad} \cdot \Delta t_{dj}^{railroad} \tag{6}$$

where i = an index for passing time of trains; j = an index for type of train class and transportation; $n_{P_{ij}}^{railroad}$ = number of passengers in railroad car; $T_{ij}^{railroad}$ = Average Daily Traffic Volume (ADTV); $u_{ij}^{railroad}$ = average unit value of time per the user; $\Delta t_{d0}^{railroad}$ = the additional delay time.

3.2.2 *Socio-economic cost model*

Indirect socio-economic losses are result of multiplier or ripple effect on economy caused by functional failure of a structure. Recently, Lee (1996) developed this functional failure cost for a specific building model as the first round and second round loss based on the I-O model. In the case of building structures, indirect socio-economic losses are relevant only to labor force of workers who are resident of the building. However, in the case of bridges, these indirect losses are influenced by not only the road user of a functionally failed bridge but also all the road users and the regional industrial sectors within the traffic networks where the bridge is located. Based on the previous study (Lee, 1996), an improved cost model was proposed, which can be reasonably applied to a bridge for the assessment of the indirect socio-economic losses incorporating the effect of traffic network (Cho et al, 2004).

The first-round losses, C_{B1}^{loss} of the indirect socio-economic losses due to functional failure of a bridge during rehabilitation activity may be obtained as

$$C_{B1}^{loss} = \sum_{i=1}^{I} \sum_{j=1}^{J} \varepsilon_{ij} Y_{ij}^{loss} = \sum_{i=1}^{I} \sum_{j=1}^{J} \varepsilon_{ij} \left[\left(\frac{\sum_{k=1}^{K} T_{ik} \cdot n_{p_{ik}} \cdot \Delta t_{d_i}}{\sum_{i=1}^{I} \sum_{k=1}^{K} T_{ik} \cdot n_{p_{ik}}} \right) \cdot \left(\frac{1}{t_{IO}} \right) \cdot v_{ij} \right] \cdot Y_{ij}^{P} \tag{7}$$

where i = an index for route in network; j = an index for sector of I-O table; k = an index for vehicle type such as owner car, taxi, truck, bus etc.; ε = economic surplus per

unit total output of production; Y^p = total output without any rehabilitation activity; t_{IO} = time interval of the I-O table; ν = sectoral participation factor, i.e. the ratio of the number of workers; T = number of average daily traffic volume; n_p = rate of passengers in vehicle; and Δt = time delay.

As the consequence of the first-round losses, the loss of capacity in one industry would likely reduce the productivity of other industries which obtain input from the industry. Accordingly, this reverberation would lead to the second round losses, C_{B2}^{loss} as follow:

$$C_{B2}^{loss} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\varepsilon_{ij} (Y_{ij}^{*} - Y_{ij}^{d}) \right] \cdot Y_{ij}^{p} \tag{8-a}$$

$$Y^d = (I - A^*)^{-1} D^* \tag{8-b}$$

where Y^* = change in gross output, estimated as $Y^P - Y^{loss}$; Y^d = new level of output; D^* = new level of demand, estimated $D^* = (I - A^*)Y^*$; A^* = the input coefficient matrix during the rehabilitation activity economy whose element is equal to $(Y^*/Y^P)A$; and A = the input coefficient matrix of the I-O model

It may be noted that the indirect cost of a bridge structure may be rationally evaluated by using Eqs. $5 \sim 8$. However, because for the assessment of road user cost and indirect socio-economic losses using the proposed cost models, highly complicated site-specific data are required, it is extremely difficult or even impossible to apply these models to each bridge. Therefore, instead of theoretical modes, an approximate but reasonable approaches were suggested by using regression analysis that utilizes site-categorizing data for each major parameter of these cost models for the practical implementation of the indirect cost. Moreover, in practice, socioeconomic losses could be approximately evaluated as a percentage fraction ranging from 50% to 150% of road user costs.

Also, it may be noticed that in the case of the PLCCA for the evaluation of design-alternatives, socio-economic losses don't have to be considered because at the design stage the critical rehabilitation or retrofit due to the failure or damage of a infrastructures are not usually required under design loading environment. However, when the reliability/risk-based LCC-effective design is to be used, the indirect cost must be considered in the assessment of the expected LCC.

3.3 *LCC models for optimum design*

Design of an infrastructure may be based on a comparison of risks (costs) with benefits. An optimal design solution, chosen from multiple alternative designs, can then be found by minimizing the LCC. Such a decision analysis can be formulated in a number of ways by considering various costs or benefits, which can also be referred to as a whole life costing or lifetime cost–benefit or cost–benefit–risk analysis. And thus the LCC may be used to assess the cost-effectiveness of design decisions. If the benefits of each alternative are the same, then the total expected LCC up to the life span T_L of an infrastructure may be formulated as follows:

$$E[C_T^{PV}(\vec{X})] = C_I^{PV}(\vec{X}) + E[C_M^{PV}(\vec{X})] + E[C_F^{PV}(\vec{X})] \tag{9-a}$$

$$E[C_M^{PV}(\vec{X})] = E[C_{RM}^{PV}(\vec{X})] \tag{9-b}$$

where $E[C_F^{PV}(\vec{X})]$ = expected (failure) rehabilitation cost for failure limit states considered in the model.

Similar to the design cost and load testing cost, the expected maintenance costs have also been estimated as a function of the construction cost (De Brito and Branco, 1995; Leeming, 1993; Piringer, 1993). However, if the suggestions are followed, then the more cost would be required as the construction cost increases, which may result in a contradiction to the assumption. Meanwhile, Wen and Kang (1997) assumed that, in their LCC analysis, the dependence of maintenance cost on the design variables under consideration would be generally weak, and thus they did not consider the maintenance cost in their LCC model. However, since the maintenance costs over a life span may be high, depending upon the quality of the design and construction of bridge under environmental stressors, the costs must be carefully regarded as one of the important costs in the evaluation of the LCC. In general, the maintenance costs can be classified in terms of the costs which are related to design variable (*e.g.*, painting cost, pavement, and etc.) and bridge management (*e.g.*, periodic routine maintenance, inspection/diagnosis). Since, in the LCC optimum design, the latter might have a minor influence on the optimum solution, expected maintenance can be ignored in the LCC formulation (Melchers 1987). Thus, in this paper, only the expected maintenance costs which are related to design variables are considered as Eq(9-b).

In case of LCC-effective optimum design based on time-variant reliability, Eq(4-a) requires the assessment of the updated annual probability of failure for possible limit states considered in the cost model such as strength, fatigue, vibration, stability, serviceability, corrosion, etc. and that of associated costs as well. Since costs may occur at different times, it is necessary for all costs to be discounted using a discount rate q. And then the general formulation for LCC-effective optimum design of an infrastructure may be represented as follows:

$$\text{Find} \quad \vec{X} \tag{10-a}$$

$$\text{Minimize} \quad E[C_T^{PV}(\vec{X})] \tag{10-b}$$

$$\text{Subject to} \quad G_j(\cdot) \le 0 \quad j = 1, 2, \ldots, N_S \tag{10-c}$$

$$P_{F_i}(\vec{X}) \le P_{F_{allow}} \quad i = 1, 2, \ldots, L \tag{10-d}$$

$$\vec{X}^L \le \vec{X} \le \vec{X}^U \tag{10-e}$$

where $\vec{X}$ = the vector of design variables; $G_J(\cdot)$ = j-th constraint (i.e., allowable stresses or resistance, combined stress/strength limits, geometric limits, etc.); N_S = the number of constraints; $P_{F_i}(\vec{X})$ = annual probability of failure for limit state i considered in the design; $P_{F_{allow}}$ = allowable probability of failure for the limit state i; and $\vec{X}^L$, $\vec{X}^U$ = the lower and upper bounds, respectively.

Also, in case of LCC-effective optimum design based on time-invariant reliability analysis, Eq(4-a) may be changed as follows:

$$E[C_F^{PV}(\vec{X})] = \sum_{i=1}^{L} [(C_{DR}(\vec{X}) + C_{IR}(\vec{X})) \cdot P_{F_i}(\vec{X})] \tag{11}$$

where P_{F_i} is the probability of failure for limit state l over design life or service life.

Expected failure cost in design life or service life, Eq (11), could be used practically and approximately for time invariant optimum design under design loadings except for the cases where time-variant extreme loading like typhoon, flooding, earthquake or other natural hazards, should be considered for more realistic LCC-effective design.

4 Practical methodologies for LCC analysis and design

Civil infrastructures such as buildings, bridges, tunnels, dams, underground structures, etc. are invariably subject to appreciable diverse uncertainties in the evaluation of time variant performance degradations and loadings and environmental effects throughout the life span of the facilities. Moreover, in normal and abnormal service conditions and operational environments, most of civil infrastructures are subject to aging, deterioration, damages due to various man-made or natural hazard, or physical process or chemical attacks. Also, the combined effect of environmental stressors, aggressive service conditions, and the absence of effective maintenance may accelerate the deterioration and/or damage failures of infrastructures. Accordingly, the evaluation of life cycle cost inevitably involves the assessment of time variant performance (condition states, carrying capacity, reliability, etc) of elements, component subsystem or system itself, associated with expected costs of maintenances, repairs, rehabilitation and retrofit, which in turn involves the assessment of expected failure (damage) probability, time variant reliability and risk of facilities. Accordingly, the LCC analysis and design itself theoretically becomes extremely complex stochastic time-variant problem, and moreover when the relevant databases are not available, the elaborate methodologies for the LCC become meaningless to come up with optimal decisions on engineering value and economics problems.

4.1 *Assessment of performance degradation*

In most conventional approaches developed so far, maintenance scenarios are based on time-controlled maintenance (TCM) interventions for which the random application time is given as input data. Since the application time of maintenance interventions is decided only based on engineering experience or statistical database, these interventions depends on neither the condition state of the structure nor the target performance level. On the contrary, performance-controlled interventions (PCM) represent maintenance actions applied when a specific condition is reached. The application times of these interventions depend on both the target performance level and previous maintenance interventions. Since the performance-controlled maintenance scenario can reflect previous maintenance intervention, economical maintenance scenario that ensures a specified structural performance can be established (Kong & Frangopol, 2004).

In Korea, various FMSs (Facility Management Systems), such as Korea BMS (KRTA, 1998), KRIFIS(KORAIL, 2003) and FMS (KISTEC, 1995), have been used only for

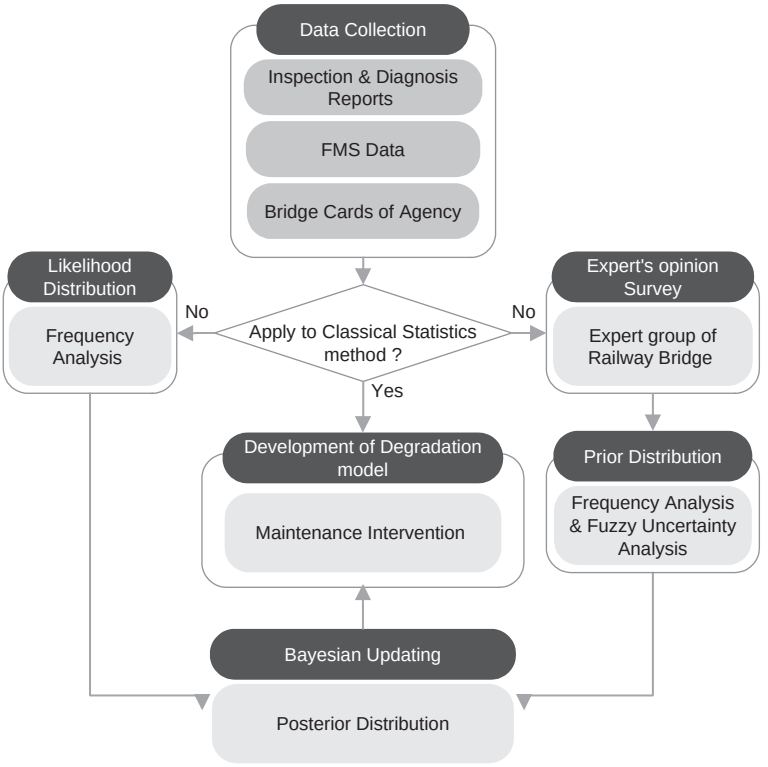


Figure 2 Procedure for development of degradation model.

conventional maintenance & management of civil infra-facilities. Most of these FMS are only built upon discrete condition states based on subjective visual inspection and NDT data.

Accordingly, in the case of developing countries like Korea, the assessment of performance degradation in terms of condition state and reliability index based on only raw database such as FMS/BMS is almost impossible because of the absence of relevant data and information on the structural performance of most infrafacilities. However, in PLCCA that includes the maintenance & repair costs associated with performance-controlled interventions at design stage, the condition deterioration may be reasonably used as performance degradation profile for the PCM-based LCC evaluation of design-alternatives because the critical rehabilitation or retrofit due to structural failure or damage are not required under design loading environments. Therefore, for PLCCA incorporating performance degradation, in terms of condition indexes at both element and system levels, may be used to construct condition deterioration profiles at design stage. However, even at design stage, a simple reliability-index degradation profile for structural members may be used in the PCM-based assessment of maintenance & repair time and cost if the statistical data of resistance deterioration of those members are available or estimated subjectively by the experts.

As an illustrative example, in a research project on the development of LCCA system for rail-road bridges (KRNA, 2007), performance degradation models in terms of condition state index for components/systems of railroad bridges have been developed by following the procedure shown in Fig. 2., based on uncertainty assessment and subsequently, performance-controlled intervention strategy that makes use of performance degradation model in terms of condition state index, which can be used to estimate lifetime maintenance probability and strategy.

The information required in the evaluation of the condition state degradation can be obtained from the statistical analysis and/or Bayesian updating using insufficient statistical data available and/or expert's opinion survey data. Detailed methods and procedures are not presented herein but referred to the reference (Cho, 2008).

4.2 *Uncertainty assessment*

Uncertainties in LCC analysis may easily lead to making the wrong decision, especially in selecting the best alternative from pairs of closely ranked competing strategies. When the uncertainties, for LCC analysis and design have to be assessed, in general as stated before, the uncertainty control and estimation independently considers three different cases corresponding to all of the three possible approaches in the following ways: 1) When historical data is sufficient and available, the probability of each variable can be evaluated by using a simple frequency analysis. 2) On the other hand, if the data is insufficient, engineering probability theories such as MCS or Bayesian approach could be used as uncertainty assessment methods. In other words, randomness and incompleteness of insufficient data is complemented by simulated or updated data. 3) Also, if the data is not available at all, the occurrence probability of each variable may have to be assessed by subjective judgments based on experienced experts' opinion and knowledge. In this case, verbal descriptive (fuzzy) uncertainty assessment has also been proven as a valuable tool for handling uncertainties due to subjective estimates (Cho et. al, 2002).

However, it may be noted that all the three approaches aforementioned involve epistemic uncertainties, feature of incompleteness, fuzziness, and biases. Therefore, for evaluating subjective judgments based on experts' experiences and knowledge, it is to be noted that the biases in expert's opinion have to be considered rationally in the uncertainty control and estimation. In general, there are two major categories of the biases -namely, motivational bias and cognitive bias. Motivational biases arise when one or more experts provide intentionally biased opinions. A primary source of this type of bias is real or perceived conflict of interest on the part of one or more experts. Cognitive biases are totally unintentional. Cognitive biases associated with epistemic uncertainties depend on the way the expert consolidates different pieces of information in formulating an opinion.

Both motivational and cognitive biases can be reduced from the aggregated results of the survey during the analysis and aggregation of the expert' estimates. However, if only few experts' opinions are available, it is important to consider the uncertainty range that represents the degree of uncertainties involved in both probabilistic parameter estimates and subjective judgments. For example, since probabilistic parameter estimates are usually based on historical records, the uncertainties are considerably affected by the use of inadequate statistical analysis methods or the lack of data for

accurate analysis. For uncertainties of subjective judgments, the major factors such as the complexity of working and health conditions for judgments and the level of education, assurance, and experience have influence on the epistemic uncertainties. Although there are a large number of factors that can be affected by the degree of uncertainties, these major factors must be considered in the uncertainty assessment because of their major contributions to uncertainties. Moreover, it is especially more reasonable to use the uncertainty range in case when the experts' opinion may not be reliable or be in error (Cho et. al., 2002).

4.3 *Optimization algorithms and procedures*

Various optimization techniques are applied to numerous optimum design and optimal decision-making problems for the selection of best alternatives with objectives of minimizing cost and/or maximizing performance. Recently, it is often used to make LCC-effective optimum design of civil infrastructures.

Recently, general optimization algorithms for the reliability-based LCC-effective optimum design of bridges were developed by the author, which consists of a structural analysis module, reliability analysis module, optimization module, and LCC evaluation module.

In reality, since an infrastructure has a large number of design variables and shows complex structural behavior, it would be impractical to directly use the algorithm for its optimum design, especially for the LCC optimization problem. Thus, various approximation techniques, structural reanalysis techniques and multi-level optimization technique as well, can be applied for numerical efficiency. Also, for some optimization problems the augmented Lagrange multiplier (ALM) method is applied, with unconstrained minimization and search techniques (Cho et. al, 1999). In particular, since bridge structures have a large number of design variables and shows complex structural behaviors, it would be impractical to directly use conventional optimization algorithm for its optimum design, especially for the LCC optimization problem. Thus, a multi-level (M/L) optimization design algorithm was introduced to improve the computational efficiency. An automatic differentiation (AD) is also incorporated into the M/L Optimization algorithm for the effective reanalysis of the main structures.

However, recently, genetic algorithms (GA) are more often used to optimize discrete design of elements or members for civil infrastructures at design stage, or to generate optimal LCC-effective maintenance scenario in the form of a multi-objective combinatorial optimization problem at maintenance stage. (Cho et. al, 1999, Cho et. al, 2001, Cho et. al, 2004, Lee et. al, 2006; Cho et. al, 2006).

4.4 *Time-variant reliability analysis*

The evaluation of expected rehabilitation cost utilizing Eq. (4) requires the reliability analysis for all failure limit states in the cost model. The reliability analysis is concerned with the calculation and prediction of the exceedance probability for a limit state in the model. Realistic limit state models have to be defined for each failure limit state. This might include critical limit state such as bending strength, shear strength, serviceability, fatigue, and so on. The cumulative life-time probability of failure can be obtained by combining the extreme live load model, the time-variant resistance model and the limit state models. And the reliability of structures is only valid for a specific point in time.

The maximum value of the live load is expected to increase over time, and structures deteriorate through aging, increased use, and specific degradation mechanisms such as fatigue and corrosion, etc. Accordingly, time-variant reliability analysis is performed to estimate annual probabilities of failure; an annual probability of failure $P_{F_i}(\vec{X}, t|T)$ of structure is used to evaluate the expected failure (rehabilitation) costs.

The probability that a structure will fail in t subsequent years, given that it has survived T years of loads, can be expressed as follows (Stewart & Hossain, 2001):

$$p_{F_i}(\vec{X}, t|T) = \frac{p_f(\vec{X}, T+t) - p_f(\vec{X}, T)}{1 - p_f(\vec{X}, T)} \tag{12}$$

where $p_f(\vec{X}, T+t)$, $p_f(\vec{X}, T)$ = the cumulative probability of failure of survice proven structures anytime during this time interval, respectively.

In case of highway bridges, the live load model used in this study (Nowak 1993) accounts for the increased effects of maximum shear and moment as more trucks pass over the bridge. This model is based on statistics of extreme values where the probability of encountering a heavy truck at the extreme upper tail of the distribution increases as the number of trucks passing over the bridge increases. The live load model uses the actual daily truck traffic and the actual bridge span lengths to predict the maximum moment and shear effects over time in a bridge.

Also, in case of steel bridges, except for high performance steel such as anti-corrosion weathering steel, almost all structural steel are subject to corrosion to some degree. As long as structural steel bridges are concerned, corrosion reduces the original thickness of the webs and flanges of the steel girders. Severity of corrosion, in general, depends on the type of steel, local corrosion environment affecting steel condition including temperature and relative humidity, and exposure conditions (Albrecht, 1983). Based on the previous study (Ellingwood et. al., 1999), a modified corrosion propagation model which considers the influence of periodic repainting effect on the corrosion process is introduced as follows:

$$p_i(t) = p_{i-1}(i \cdot T_{REP}) + C \cdot (t - i \cdot T_{REP} - T_{CI})^m 1 \quad 1 \text{ for } \quad (i) \cdot T_{REP} + T_{CI} \leq t < (i+1) \cdot T_{REP} \tag{13-a}$$

$$p_i(t) = p_{i-1}(i \cdot T_{REP}) \text{ otherwise} \tag{13-b}$$

where $p_i(t)$ = corrosion propagation depth in μm at time t years during i-th repainting period; C = random corrosion rate parameter; m = random time-order parameter; and T_{CI}, T_{REP} = random corrosion initiation and periodic repainting period, respectively.

On the other hand, more theoretical and elaborate time-variant reliability analysis models such as lifetime seismic reliability analysis of deteriorating bridges for LCC-effective seismic design are not discussed herein because of limited space. But the methodology for lifetime seismic reliability analysis basically involves the following analysis: i) Probabilistic Lifetime Seismic Response Analysis (PLSRA); ii) Probabilistic Lifetime Seismic Damage Analysis (PLSDA) using the results of PLSRA; iii) Probabilistic Seismic Hazard Analysis (PSHA); and iv) Lifetime Seismic Reliability Analysis

(LSRA) using the convolution of seismic hazards with the probabilities of structural damage (Cho et. al, 2006).

5 LCCA systems

5.1 *LCC system model and knowledge-based data base*

Though a number of LCC systems were developed so far, there are only few systems practically applicable to the real problems in practice. Moreover, nowadays practicing engineers are always concerned about the availability of integrated system applicable for LCC-effective decisions on design and rehabilitation of various kinds of infrastructures. For instance, in tender design phase for bids, engineers need some powerful LCC assessment tools for the selection of construction type, durable construction materials, and construction methods. For these problems, the PLCCA model incorporating performance degradation model may be applied more effectively in practice.

From this point of view, a LCC system code practically applicable in practice was developed by the author with other research associates mainly based on PLCC. The project was funded by the Ministry of Construction and Transportation (MOCT), which was carried out as one of the most strategic, important SOC research projects. Fig. 3. shows the schematic diagram for integrated LCC system model mainly for bridge structures (KICT, 2006; Lee et. al, 2004).

First of all, it may be noted that the construction of the knowledge-based DB is simply the most important part in the development of the LCC software system. Fig. 4. shows the conceptual diagram for the knowledge-based DB. As shown in the figure, the main function of the knowledge-based DB is to store and assess all the cost and uncertainty data as well as all the information, such as essential information on rehabilitations (i.e., rehabilitation cost, time from first rehabilitation and period of subsequent rehabilitations corresponding to maintenance strategies, increase of performance or decrease of deterioration rate after rehabilitation, duration of rehabilitation activity, work-zone condition, etc.), site-specific information on site characteristics and traffic network, etc., and various information associated with indirect cost assessment. The information can be acquired by historical data, expert's opinion, engineering practice and analytical damage prediction model which comprise of key components of the DB. Accordingly, these cost data and information are made to be stored and assessed in the knowledge-based DB of the LCCA systems. The input data related to the type of damages and the type of maintenance methods are incorporated into data base based on the sources available in Korea, such as "Method for Extension of Service life of Road Bridges" (KISTEC, 2000) and "Data for Design Guide to Maintenance and Repair of Infrastructure" (OCFM, 2002), etc. Also, the data related to the time from first rehabilitation and period of subsequent rehabilitations with respect to each alternative maintenance strategies associated with 'do nothing', 'essential maintenance', 'preventive maintenance' are based on the survey of expert's opinion and a reference such as "An Enhanced Strategy for Infra Facilities Safety Management System based on LCC Concept" (KICT & KISTEC, 2001). In addition, the information on evaluation of the indirect user cost and socio-economic losses (i.e., Input-Output table, the ratio of the number of workers, etc.) are also included in the database.

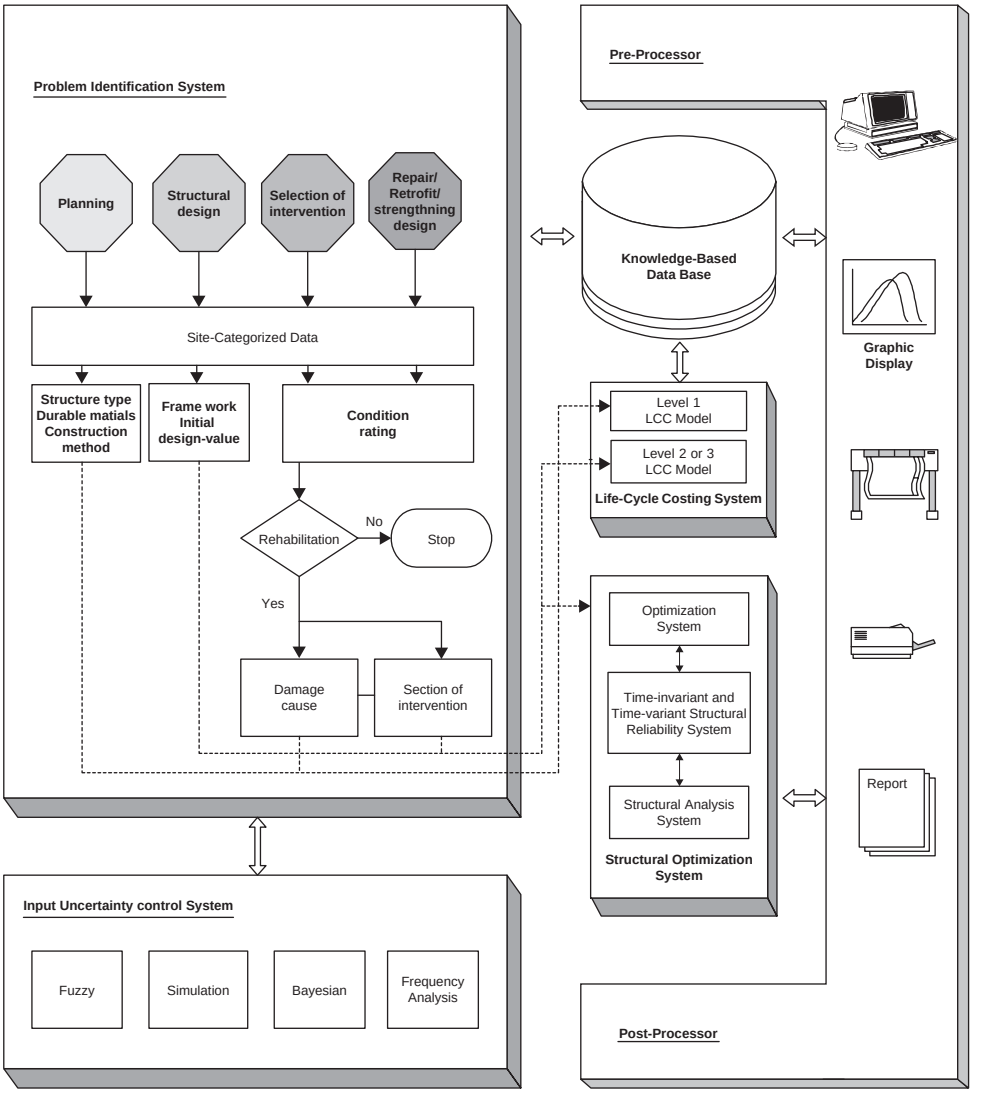


Figure 3 Schematic integrated LCC system model.

5.2 *LCCA system for design stage*

Recently in Korea, a number of LCCA systems have been developed and some of them are still under development by the author with other research associates. Among these systems, for example, the first standard LCCA system for highway steel bridges called LCCSTEB was developed using PLCCA model as a software package on personal computers with MS-windows platform and especially designed for individual steel bridges (KICT, 2006). It is used to make a decision for the selection of construction

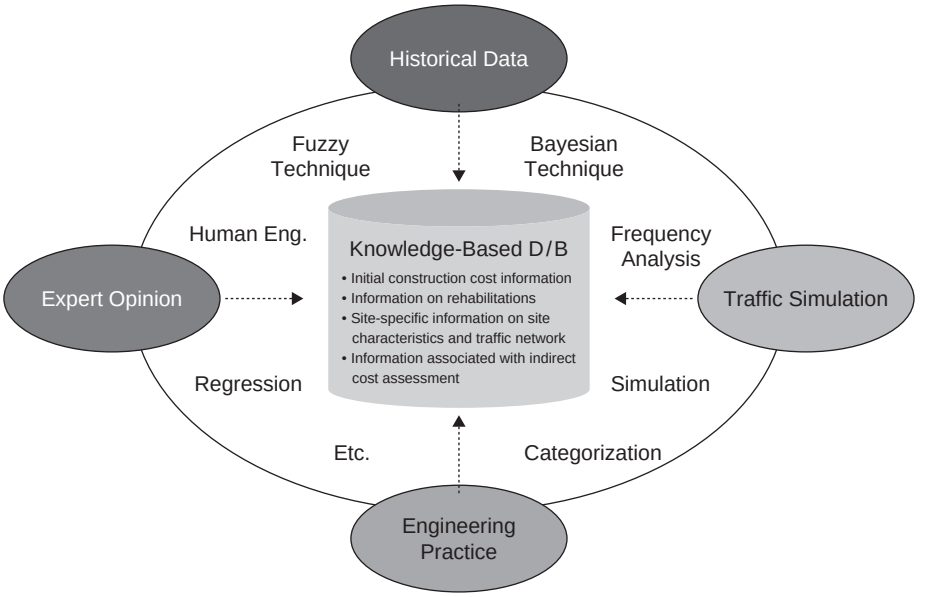


Figure 4 Conceptual diagram for knowledge-based data base.

type, durable construction materials, and construction methods, etc. Also, it will have additional ability to evaluate appropriate points in time of maintenance interventions of steel bridges associated with performance deterioration in the future. Basic theory of REHAB developed by Kong (2001) may be used for the additional development. It consists of the LCC computing and the DB management systems. Input data is automatically generated by the DB management system in LCCSTEB. Fig. 5. shows the simplified flow chart and a screen snapshot of LCCSTEB. Major sub-modules and their functional descriptions are presented in the flow chart. The followings are brief descriptions of development process for the DB management system.

It may be stated that DB has to be continuously complemented and updated for using LCCSTEB efficiently during lifetime of bridges. Therefore, a DB management system will be needed to provide user convenience and automatically updated data to LCCSTEB through DB processing and analyzing.

To build a DB for analyzing LCC of steel bridges, integration and networking of between DB systems in Korea were performed. As a result, a knowledge-based DB system has been developed for computing and estimating practical life-cycle cost at construction stages.

Raw data is not particularly useful in LCCA. The DB management system has an important function of generating input data into LCCSTEB with the refined DB. The refined DB consists of maintenance time data and cost data obtained by statistical analyses using raw DBs. Those raw DBs were obtained from existing bridge management systems (BMS), surveys, and so on as shown in Fig. 6. This refined DB is the core of the knowledge-based DB system. Main functions of DB management system are to

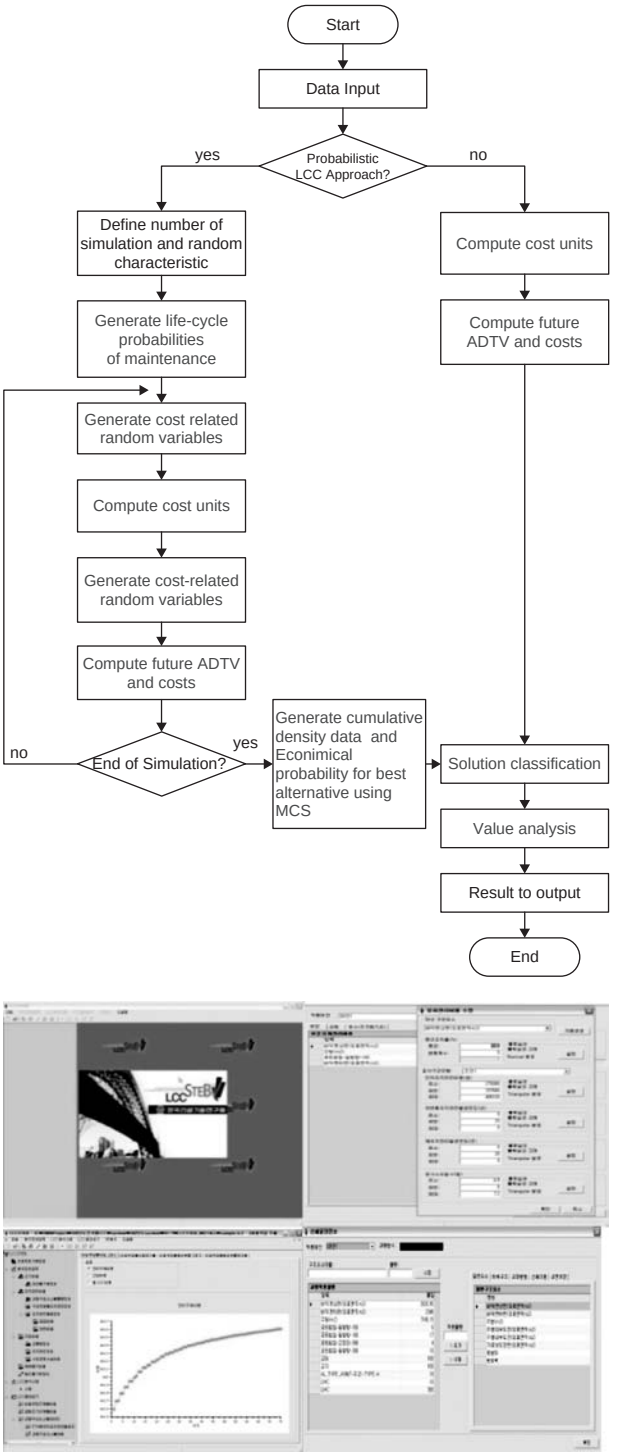


Figure 5 Life cycle cost analysis algorithm and a screen snapshot of LCCSTEB (at design stage).

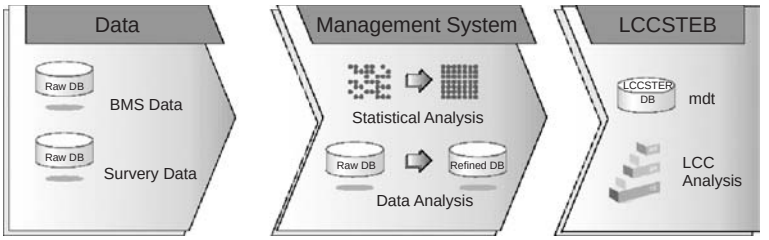


Figure 6 Data flow in LCCSTEB.

manage data, to analyze data and to transfer date into LCCSTEB. The DB management module manages BMS data, survey data, field survey data, and reports data in the form of file. The DB analyzing module analyzes gathered DBs, and the DB transferring module transfers input data from the DB analyzing module into LCCSTEB.

Also, another LCCA system for railroad bridges, called RailBLCC has been also developed by the author with other research associates, which also used PLCCA model and evaluates the expected maintenance costs based on performance-degradation regression model. Statistical and regression uncertainties are estimated and updated based on historical FMS/KRIFIS data and expert's opinion survey, which are used to define a maintenance intervention because maintenance characteristic and environment of railroad bridges are different from those of highway bridges (KRNA, 2007).

6 Application examples

6.1 *Application of LCC models for design-alternative analysis & optimum design*

The following is a typical example for the selection of LCC-effective optimal bridge type for bids at design stage, and the associated example for LCC-effective optimal design of the selected bridge (Cho, 2002). In this example, four bridge types – namely, Steel Box Girder bridge, Four Main Plate Girder bridge, PSC-Box Birder bridge by Free Cantilever Method (FCM), PSC-Box Girder bridge by Full Staging Method (FSM) are considered as the alternatives. The bridge will be constructed as a part of a rural highway which has relatively moderate ADTV, expected to increase from 21,000 to 46,000 over 75-year design life. At the conceptual design stage, optimal bridge type should be selected by considering Value Engineering (VE) aspects incorporating economy (LCC), safety, aesthetic, workability, functionality, maintenance convenience and environmental effect, etc. Thus, for this bridge design problem, the LCCA is carried out as a part of VE evaluation of the alternative bridge types. In evaluating the LCC for each alternative, the maintenance is assumed to be conducted by preventive maintenance, which is based on the survey of expert's opinion (KIST and KISTEC, 2001). The more detailed data for each alternative applied in LCCA can be found in the reference (KECC — Korea Engineering Consultant Corp., 2002). The LCCA for this example is also performed by using PROLCC and LCCSTEB. The same data for cost parameters presented in Table 1. are used for the LCCA of this example.

Table 1 Major parameters used in design-alternative analysis and optimum design.

Parameter	*Value*	*Unit*	*Reference*
The traffic accident cost	0.12	bil. won	KOTI (http://traffic.metro.seoul.kr)
traffic accident rate during repair work activity	2.2	mil. vehicle/km	
traffic accident rate during normal condition	1.9	mil. vehicle/km	
The value of fatality	3.5	bil. won	Lee and Shim (1997)
The value of injury	21	mil. won	
The hourly driver cost	21,517	won/person	KOTI
Discount rate	4.51	%	KISTEC (2000)

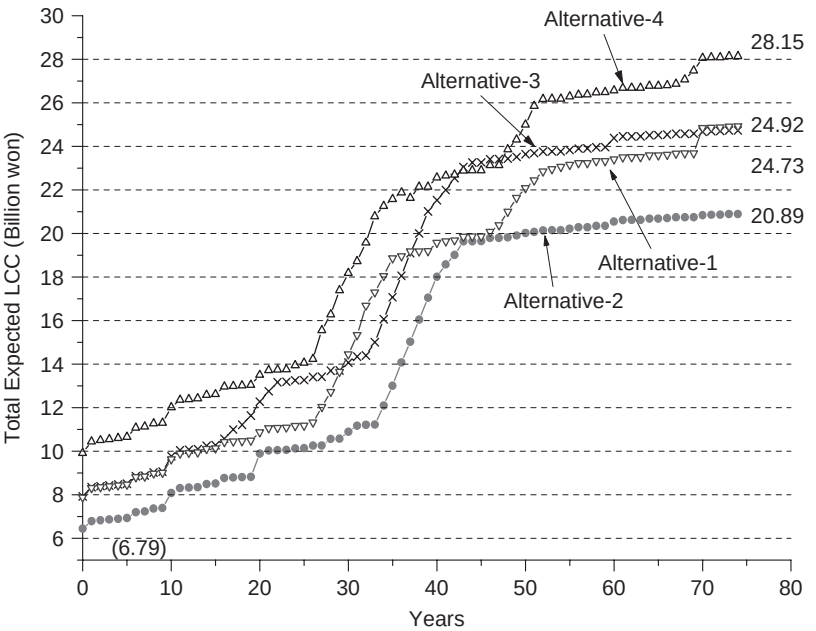


Figure 7 Total expected LCC for each alternative.

In Fig. 7., the results of LCCA for the each alternative are comparatively shown. As shown in the figure, the expected LCC of the alternative-2 — namely, Four Main Plate Girder bridge — shows the most economical result estimated as 20.89 billion won. Also, Fig. 8. is cumulative distribution function for the each alternative, respectively.

Fig. 8. shows that expected LCC of alternatives at 90% confidence level are 29, 26, 29 and 33 billion won, respectively; alternative 2 is most economical among others like a DLCCA. But PLCCA provides more information in terms of percentile probability (occurrence possibility %) of one alternative relative to other alternatives. Thus, in view of decision-making, PLCCA is a lot more useful than DLCCA.

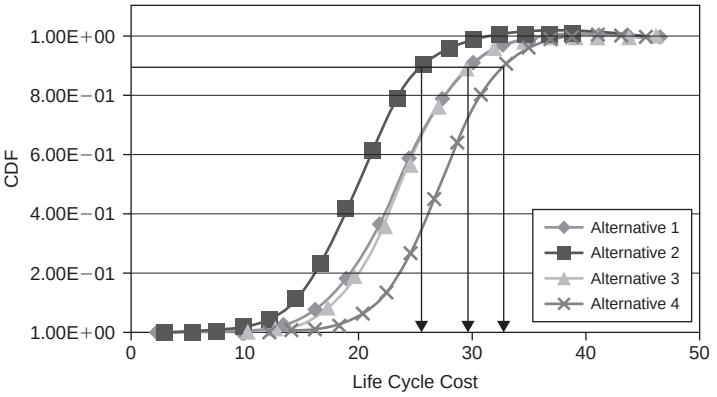
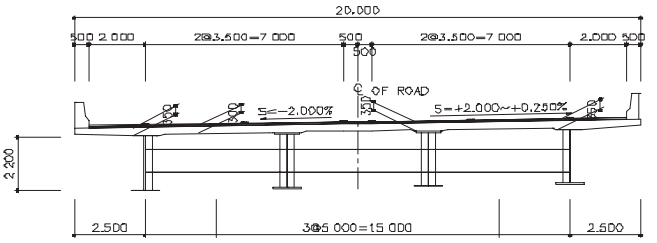
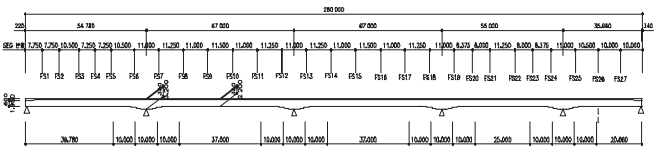


Figure 8 CDF of LCC for each alternative.



(a) Typical section



(b) Profile

Figure 9 Typical section and profile of four main plate girder bridge.

According to the results of VE evaluation incorporating those results of LCCA, the optimum alternative is chosen as the alternative-2, which is the same as that of the LCC assessment. As a result, the optimal bridge, Four Main Girder bridge, is shown in Fig. 9., whose roadway is 280 m long and 20 m wide, and has two lanes for each way. In the LCC optimization for the alternative-2, the design variables are selected as upper/lower flange and web thickness, and girder height. The design constraints are formulated based on Korean Bridge Design Specification.

The problem of optimum LCC design of the bridge is formulated as that of minimization of the expected total LCC that consists of initial cost, expected (failure) rehabilitation cost for strength and serviceability limit states. For instance, the cost associated with expected strength failure can be obtained from the probability of

Table 2 Maintenance interventions.

Element	*Maintenance Intervention*	*Unit Cost (Thousand Won)*	*Application Rate (%)*	*Improved Condition Index*
Steel Girder	Painting (PG)	30.0/m^2	100%	30%
	Welding (WE)	20.0/m	10%	40%
	Bolting (BO)	3.0/ea.	10%	30%
	Steel Attachment (SA)	750.0/m^2	10%	70%
	Replacement (RG)	2,000.0/m^2	100%	100%
Slab	Epoxy Injection (EI)	3.0/m	30%	24%
	Waterproofing (WP)	27.6/m^2	80%	24%
	FRP Attaching (FA)	78.9/m^2	50%	72%
	Replacement (RS)	150.0/m^2	100%	100%
Pavement	Surface Treatment (ST)	10.0/m^2	10%	30%
	Cutting-overlay (CO)	20.0/m^2	30%	80%
	Patching (PA)	10.0/m^2	5%	40%
	Re-pavement (RP)	25.0/m^2	100%	100%

strength failure and the damage cost that consists of strengthening cost, human and property losses costs, road user costs due to traffic closure, and indirect regional economic losses.

The strength failure limit-state in terms of bending and shear of a main girder is defined as the ultimate limit state failure of flange and web of main girder system. As shown in Fig. 7., the initial cost for initially assumed sections of the bridge is 6.79 billion won. The steel plate adhesion method is assumed to be used for the strengthening of main girder, the estimation criteria for unit repair cost is taken as 266.75 thousand won/m^2, based on the reference (KISTEC, 2000). Major parameters related to human and property losses costs and road user costs are given in Table 2. Though, detailed information on the assessment of indirect regional economic losses, uncertainty data related to reliability analysis and detailed procedure are not presented in this paper, due to space limitation which can be found in the references (Cho et. al, 2001).

In this paper, only the economical aspects are focused and discussed. Since the LCC optimum design can be equivalently achieved by the initial cost optimization with a reasonable allowable stress ratio (Cho, 2001a), the initial cost optimization is carried out by varying allowable stress ratio from 75% to 100% with increment of 5%.

As shown in Fig. 10., the total expected LCC's for the 6 different levels of allowable stress ratio are estimated based on the Initial Cost Optimization (ICO).

It can be seen that the Initial Cost Optimization at the allowable stress ratio of 80% equivalently achieves the near optimum LCC design point. In Fig. 11., for each of the three different optimizations — the conventional ICO, the ICO with allowable stress ratio of 80%, which is called ELCCO (Equivalent LCC optimization), and LCC Optimization (LCCO) — the results of initial cost, expected failure cost and total expected LCC are shown respectively. As shown in the figure, the results of the equivalent LCC (the ICO with optimal allowable stress ratio of 80%) and the LCCO are very similar.

However, in Fig. 11., the results for the ICO are extremely different from other cases, mainly because more slender optimum sections of the ICO results in higher

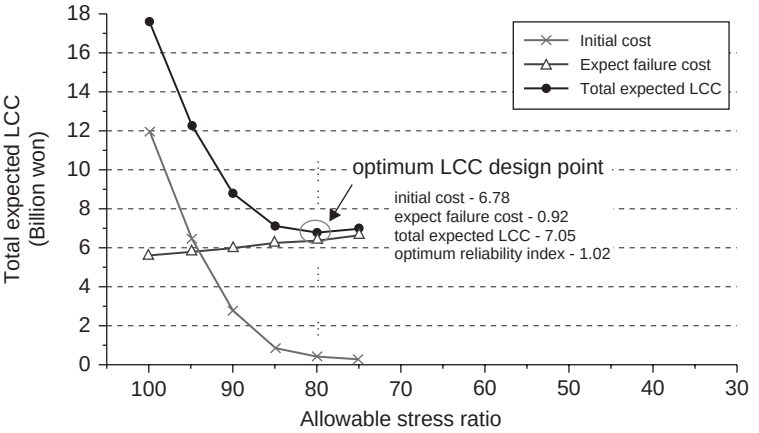


Figure 10 LCC of initial cost optimization to variations of allowable stress ratio.

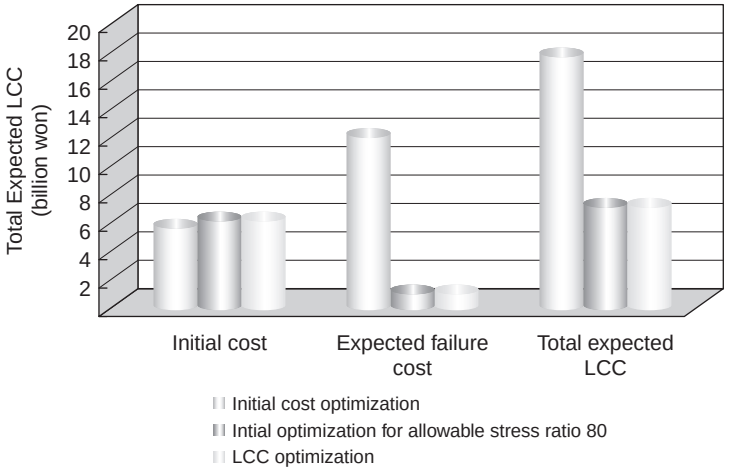


Figure 11 Comparison of costs of each optimization.

failure probably, which, in turn, yield higher expected failure costs. In the case of the ICO, the initial cost is little more economical by about 0.54 billion won compared with the case of LCCO. However, the LCCO can significantly reduce the total expected LCC by 10.47 billion won. Therefore it may be possibly stated that for the alternative-2, the more economical design may be induced in view of LCC if the initial cost optimization for allowable stress ratio of 80% or LCC optimization is applied. It has been invariably found that in a great number of other applications the equivalent LCC optimization, ELCCO, provides almost same results and thus may be effectively used instead of more rigorous and elaborate LCC optimization in practice, which means especially ELLC is very effective tool at basic design phase for LCC-effective optimum design of civil infrastructures.

Table 3 General data of steel box girder bridge.

Bridge type (Name)	*Three-span continuous steel box girder bridge (Myung-Gok Bridge)*		
Bridge length (m)	40 + 50 + 40	Type of Steel	SM 490 (fy = 320 MPa; fa = 190 MPa)
Bridge width (m)	12.145		
Skew	90°	Concrete	Strength : fck = 27 MPa
No. of lane (design)	4 (3)		Modulus ratio of elasticity : 8
No. of box	2		Reinforcing bar : SD40 (fy = 400 MPa)
Design load	HS-20		

6.2 *Application of LCC models for optimum design based on time-variant reliability analysis*

This example presents a practical and realistic LCC methodology for the LCC-effective optimum design of steel bridges considering time effect of bridge reliability under environmental stressors such as corrosion and heavy truck traffics. To demonstrate the LCC-effectiveness for optimum design of a steel bridge, a steel box girder bridge having three continuous spans (40 m + 50 m + 40 m) and a total length of 130 m is considered as a numerical example. The general data for the illustrative bridge is shown in Table 3. Bridge profile and design group, and typical section of the bridge are also shown in Fig. 12 (Cho, et. al, 2006)

As shown in Fig. 13., in the LCC optimum design of the bridge, the height of web (h_w), the thickness (t_{fu}, t_{fl}) of the upper and lower flange, the web thickness (t_w) are selected as the design variables.

It is assumed that the bridge will be constructed as a part of a typical rural highway which has relatively moderate ADTV. Fig. 14. and 15. show a highway network and modeling for traffic network analysis used for this study. And Table 4. represents the expected traffic volumes of typical rural highways for 20 years after construction.

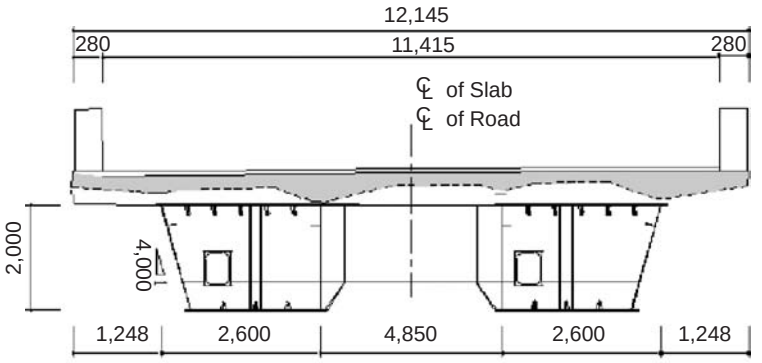
In the optimization, all the codified behavior and side constraints are considered by including various allowable stresses, combined stress limits, fatigue stress, local buckling and geometric limits based on the Korean Standard Bridge Design Code (KSBDC) of KRTA (Korea Road and Transportation Association, 2000).

Due to space limitation, detailed information for these techniques are not presented in the paper, but are referred to the reference (Cho, 2006).

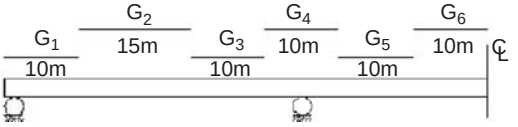
The LCC functions considered in the LCC optimization consist of initial cost, expected life-cycle maintenance cost and expected life-cycle rehabilitation costs including repair/replacement costs, loss of contents or fatality and injury losses, road user costs, and indirect socio-economic losses.

The estimation of construction unit costs as shown in Table 5 is based on the price information of the Research Institute of Industrial Science and Technology (RIST, 1998). The design cost and load testing cost are assumed as 7% and 3% of construction cost, respectively (De Brito and Branco, 1995).

Based on the various sources available, such as the CSR's price information, the opinions of the field experts who engaged in the construction, and the references available (OCFM, 2002; KISTEC, 2000; KIST and KISTEC, 2001), in order to estimate rehabilitation costs, the required data related to each failure limit state, such as



(a) Typical Section



	upper flange	web	lower flange	
G1	1.0	1.2	1.2	<- thickness (mm)
G3	1.2	1.2	1.4	(actual design size)
G3	1.6	1.2	1.6	
G4	2.6	1.2	2.4	
G5	1.6	1.2	1.6	
G6	1.2	1.2	1.4	

(b) Bridge Profile and Design Group

Figure 12 Typical section, bridge profile and design group of example bridge.

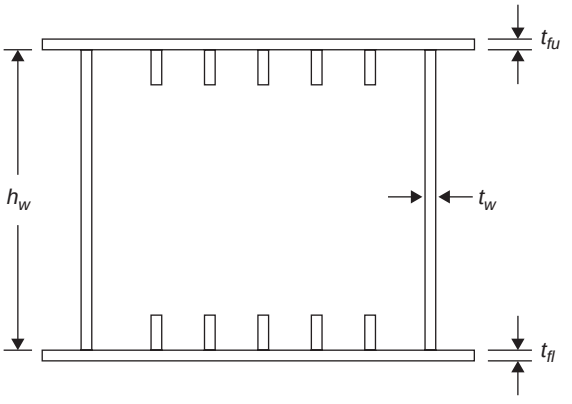


Figure 13 Design variables of steel box girder bridge.

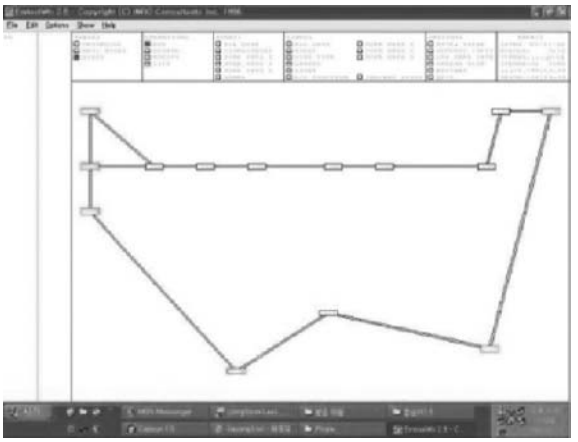


Figure 14 Highway network near the bridge site.

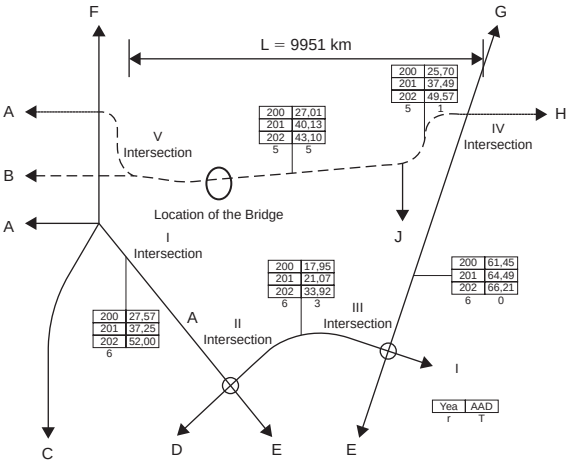


Figure 15 Modeling for traffic network analysis.

countermeasures for rehabilitation, unit direct rehabilitation cost, expected period for rehabilitation, and work-zone condition during rehabilitation activity, are summarized in Table 6.

Moreover, major parameters for the estimation of indirect rehabilitation costs of this example are same as those of Example 1 as showing Table 1.

LCC-effective optimum design of a steel bridge may be affected by local corrosion environment. Thus, in this example, to demonstrate the effect of local corrosion environment on the cost effectiveness of the LCC optimum designs of the illustrative bridge, various LCC optimizations are performed considering 5 different cases (see Table 7.). For estimating expected rehabilitation costs and the mean and C.O.V of the maximum live load moment/shear, the moderate ADTV given in Table 4. is only used in this section in order to focus on the effects of corrosion. In addition to the 5 cases, since

Table 4 Expected future ADTV of typical rural highway.

Year		*2007*	*2011*	*2015*	*2019*	*2023*	*2027*
Car	I	29,450	42,486	55,760	69,408	81,320	86,162
	II	29,762	43,774	57,432	71,480	83,750	88,734
	III	27,508	40,430	53,103	66,108	77,455	82,065
Bus	I	3,419	4,771	6,002	7,147	8,008	8,292
	II	3,522	4,916	6,180	7,360	8,246	8,543
	III	3,225	4,543	5,713	6,807	7,628	7,895
Truck	I	19,179	27,541	35,462	43,310	49,800	52,286
	II	19,757	28,376	36,525	44,599	51,289	53,848
	III	18,262	26,213	33,772	41,249	47,436	49,804

Descriptions
I :V intersection ~ location of the bridge
II : location of the bridge ~ J
III : J ~ IV intersection

Table 5 Construction unit costs.

	Structural Steel	*Weathering Steel*
Labor Cost (×10,00 won/ton)	133,09	133.09
Material Cost (×10,000 won/ton)	41.99	51.918
Paint or Repainting Cost (×10,000 won/m^2)	23.60	–

Table 6 Data related to limit states for estimating rehabilitation costs.

Limit State Countermeasures	*Strength*		*Fatigue*	
	Retrofit		*Bolting repair*	
	Structural Steel	*Weathering Steel*	*Structural Steel*	*Weathering Steel*
Unit direct rehabilitation cost (×10,000 Won)	175/ton 3.12/ton (Disposal)	185/ton 3.12/ton (Disposal)	450/location	
Expected period for rehabilitation (month)	4		3 weeks	
Work-zone condition	Partially traffic closing (2 lane)		–	

the LCC optimum design can be alternatively achieved by the initial cost optimization with a reasonable allowable stress ratio (=design stress/allowable stress) (Cho, 1998), which may be called 'Equivalent LCC optimization', the total expected LCCs for the 9 different levels of allowable stress ratio (60% to 100% with increments of 5%) are estimated based on the Equivalent LCC optimization procedure considering 3 different cases (see Table 7.).

Table 7 Cases considered in LCC optimum designs of the illustrative bridge.

Case ID	*Design Methodology*	*Type of Steel*	*Corrosion Environment*
Case CD	Conventional Design	Structural Steel	Urban corrosion Env.
Case ICO	Initial Cost Optimum Design		Urban corrosion Env.
Case O1	LCC Optimum Design		Urban corrosion Env.
Case O2			Rural corrosion Env.
Case O3		Weathering Steel	–
Case I	Equivalent LCC optimization	Structural Steel	Urban corrosion Env.
Case II			Rural corrosion Env.
Case III		Weathering Steel	–

In order to examine the relative effects of the various design cases on the LCC costs, all the costs such as the initial cost, expected maintenance cost, expected rehabilitation cost (expected retrofit cost, expected fatigue repair cost), and total expected LCC of each design case, as shown in Table 8. and Fig. 17., need to be compared with one another. It is found in the tables that the initial cost for the initial cost optimizations from Cases IOD is 0.687 billion won, while those of the LCC optimizations from O1 ~ O2 is 0.716 ~ 0.718 billion won, respectively. Therefore, it can be observed that the initial cost of the initial cost optimization is decreased by about 4.1 ~ 4.4% compared with those of the LCC optimizations. Thus it may be stated that, in general, the initial costs of the LCC optimizations may be slightly increased, as expected, compared with that of the initial cost optimizations.

Meanwhile, the total expected LCC of the optimum design from Cases IOD is 1.012 billion won. Whereas those from Cases O1 ~ O2 (the LCC optimizations) are 0.860 ~ 0.865 billion won, respectively. Therefore, it may be definitely stated that, in terms of the total expected LCC, the LCC optimizations are much more economical than the initial cost optimizations (by about 14.5 ~ 15.0%) because the failure probabilities (for strength and fatigue) in the initial cost optimizations are much higher than those in the LCC optimizations. Also, note that Table 8. and Fig. 17. clearly show the advantages of weathering steel when used for steel bridge design. Though the initial cost of the weathering steel from Case O3 is more expensive by about 3.5 ~ 3.8% compared with those of Cases O1 ~ O2, the total expected LCC from Case O3 is much more economical by about 9.0 ~ 9.6%. These trends may be evidently attributed to more expensive expected maintenance cost (repainting cost) and expected rehabilitation cost (due to the disadvantage of the structural resistance degradation as survival age increase) of conventional structural steel compared with those of weathering steel.

Also, it has been found that that the total expected LCCs from equivalent optimum LCC design (Cases I ~ III) are very close to that of the LCC optimization (Cases O1 ~ O3) only with the difference of around 0.1 ~ 0.5%. Thus, as an alternative approach to the LCC optimization, the initial cost optimization with a reasonable allowable stress ratio, i.e. the stress limit to be used for conventional design in practice, can be utilized to implicitly accomplish a near optimum LCC of steel bridges.

Fig. 16. shows the results of initial cost and total expected LCC of the initial cost optimization according to the variations of design stress level for Cases I ~ III, respectively.

Table 8 Results of optimum design of the illustrative example bridge.

Case		*Conventional Design (Case CD)*	*Initial cost Optimum Design (Case IOD)*	*LCC optimization and Equivalent LCC optimization*					
				Urban Env.		*Rural Env.*		*No Corrosion*	
				Case O1	*Case I*	*Case O2*	*Case II*	*Case O3*	*Case III*
Height of web (m)		2.00	2.09	2.50	2.50	2.50	2.50	2.50	2.50
Upper flange	Design group 1	10/12/12	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10
thickness/	Design group 2	12/12/14	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10
Web thickness/	Design group 3	16/12/16	10/10/10	11/10/10	12/10/10	11/10/10	12/10/10	11/10/10	12/10/10
Lower flange	Design group 4	26/12/24	14/10/13	19/11/18	20/11/18	19/11/18	20/11/18	19/11/18	20/11/18
Thickness (mm)	Design group 5	16/12/16	10/10/10	11/10/10	12/10/10	11/10/10	12/10/10	11/10/10	12/10/10
	Design group 6	12/12/14	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10	10/10/10
Cost (billion won)	Initial cost	0.81755	0.68687	0.71826	0.71994	0.71595	0.71994	0.74335	0.74934
	Expected maintenance cost	0.09259	0.09505	0.10611	0.11245	0.10611	0.11245	0.00000	0.00000
	Expected retrofit cost	0.01541	0.12052	0.00863	0.00801	0.00319	0.00299	0.00308	0.00210
	Expected fatigue repair cost	0.04950	0.10911	0.03233	0.02603	0.03446	0.02011	0.03565	0.02999
Total Expected LCC		0.97505	1.01155	0.86533	0.86643	0.85971	0.85549	0.78209	0.78143

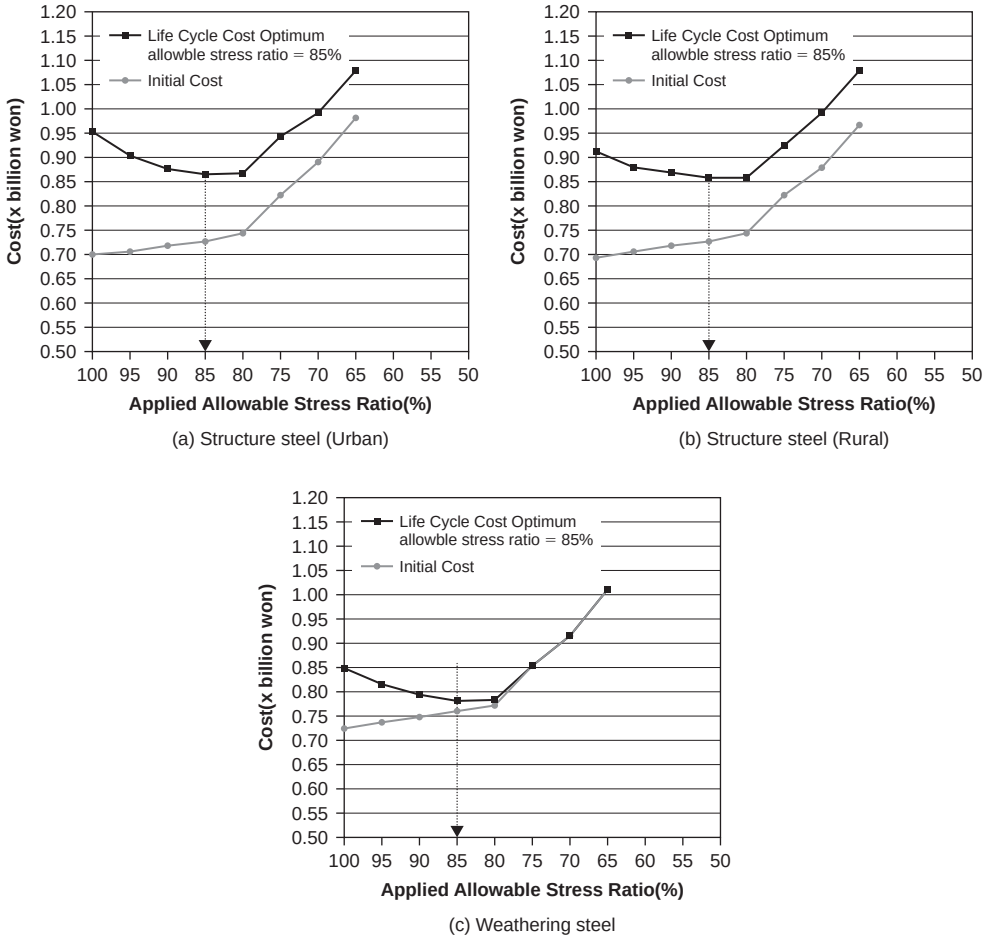


Figure 16 Initial cost and total expected LCC according to variation of design stress level.

As shown in these figures, the optimum LCC design point for Cases I, II and III are achieved similarly at the design stress with 85% of the allowable stress.

7 Concluding remarks

In the paper, a critical review is presented of the state-of-the practice and recent progress of LCC analysis and design of civil infrastructures, with emphasis on improved but practical approaches to LCC methodologies and uncertainty assessment as well. The author spotted the current problems and critical issues and then introduced practical approaches to the application of LCC analysis and design appropriate especially in developing countries like Korea.

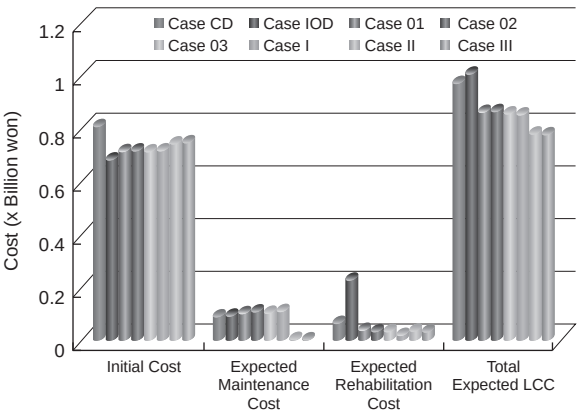


Figure 17 Comparison of costs at moderate ADTV region.

Accordingly, some important concluding remarks can be summarized as follows:

1. Practical approaches, rather than theoretical approaches, to LCC analysis and design are inevitable in practice when statistical database suitable and applicable for the LCC are not available. Moreover, the current national/public database for various infrastructures have to be completely renovated or reconstructed so as to compatible with database of LCCA system in order to implement the LCCE applicable in practice based on available database.
2. Since the current LCCA systems usually evaluate lifetime expected maintenance & repair costs of the LCC based on time-controlled interventions, the LCCA incorporating the performance degradation profile in terms of condition state index and reliability index may have to be used in practice for more realistic assessment of the LCC at design stage.
3. More upgraded reliability-based LCCA including lifetime expected (failure) rehabilitation costs against various risk of hazards may have to be used in practices especially for most of major infrastructures located in hazard prone regions where the structure are invariably subject to manmade and/or natural hazards such as explosion, flooding, seismic and earthquake, etc.
4. It has been invariably found that the equivalent LCC optimization provides almost same result and thus may be effectively used in practice instead of more rigorous and elaborate LCC optimization, mainly because LCC-effective optimum design of civil infrastructure, either time invariant or time variant, involves too much complex and elaborate procedure for lifetime reliability and LCC evaluation.
5. LCCA practice for various civil infrastructures mostly depend upon proprietary S/W's used by numerous VE/LCC consulting firms based on user supplied inputs, which are not reliable but provide only nominal results for LCC analysis and design. Therefore, the R&D for LCCE systems may have to be more actively conducted in parallel with the development of user-friendly internal knowledge-based database for each different kind of civil infrastructures.

6. So far, since the development projects for of standard guides and software systems for LCCE are limited mostly to road pavement, buildings, and bridges, the research and development of LCCE system should be extended to other major infrastructures such as tunnels, dams, underground facilities, harbor facilities, power station, and life-line systems among others.

References

Albrecht, P., S-N Fatigue Reliability Analysis of Highway Bridges, Probabilistic Fracture Mechanics and Fatigue Methods: Application for Structural Design and Maintenance, ASTM STP 798, 1983

Cho, H.N., Life cycle cost effectiveness for design and rehabilitation of civil infrastructures. *Proceedings of SEWC02*, Yokohama, Japan, 2002

Cho, H.N., Practical Approaches to the Application of Life-Cycle Civil Engineering for Infrastructures in Korea, *International Association Life-Cycle Civil Engineering'08, Keynote*, 2008

Cho, H.N., Ang, A.H.-S. & Lim, J.K., Cost-effective optimal target reliability for seismic design and upgrading of long span PC bridges. *Proceedings of the 8th International Conference on Applications of Statistics and Probability*. Newcastle, Australia: Department of Civil, Surveying and Environmental Engineering, University of Newcastle, 1999

Cho, H.N., Choi, H.H. & Kim, Y.B., A risk assessment methodology for incorporating uncertainties using fuzzy concepts. *Reliability Engineering and System Safety*. 78(2): 173–183, 2002

Cho, H.N., Kim, J.H., Choi, Y.M. & Lee, K.M., Practical Application of Life-Cycle Cost Effective Design and Rehabilitation of Civli Infrastructures. *3th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*. Lausanne, Switzerland, 2003

Cho, H.N., Lee, K.M., Cha, C.J. & Kim, J.H., Life-Cycle Cost Effective Optimum Design of Orthotropic Steel Deck Bridges under Corrosion Environment, *4th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, FL, USA., 2005

Cho, H.N., Lee, K.M., Choi, H.H. & Lim, J.K., Seismic reliability analysis of deteriorating bridges considering environmental stressors. *IABMAS. Proceedings of the 5th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, Seoul, Korea, 2006

Cho, H.N., Lee, K.M. & Choi, Y.M., Life-Cycle Cost Effective Optimum Design of Steel Bridges. *Journal of Constructional Steel Research* 60: 1585–1613, 2004

Cho, H.N., Min, D.H. & Lee, K.M., Optimum Life-Cycle Cost Design of Orthotropic Steel Deck Bridges. *Steel Structures*. 1(2): 141–152, 2001

Cho, H.N., Choi, H.H., Lee, K.M. & Park, K.H., Optimal seismic retrofit and maintenance strategy for steel bridges using Life-Cycle Cost analysis. *IABMAS. Proceedings of the 5th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, Seoul, Korea, 2006

Cho, H.N., Lee, K.M., Choi, H.H. & Lim, J.K., Seismic reliability analysis of deteriorating bridges considering environmental stressor, *IABMAS. Proceedings of the 5th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, Seoul, Korea, 2006

Ellingwood, E.R., Naus, D.j., Condition Assessment and Maintenance of Aging Structure in Critical Facilities A probabilistic Approach, Case Study in Optimal Design and Maintenance Planning of Civil Infrastructure Systems, ASCE, pp 45–56, 1999

Frangopol, D.M., Lin, K.Y. & Allen C. Estes., Life-Cycle Cost Design of Deteriorating Structures. *Journal of Structural Engineering, 1230(10)*: 1390–1401, 1997

Kong, J.S. & Frangopol, D.M., Prediction of reliability and cost profiles of deteriorating structures under time- and performance-controlled maintenance. *Journal of Structural Engineering*, 130(12): 1865–1874, 2004

Korea Infrastructure Safety and TEchnology Corporation, *Method for extension of service life of road bridges. Final report*. Ilsan, Korea Road and Transportation Association, 2000

Korean Railroad, *Korean Railroad Integrated Facilities Information System*. DJ, Korean Railroad, 2003

Korea Rail Network Authority, *Development of Life-Cycle Cost Analysis System of railroad bridges in design stage*. Daejeon, KRNA, 2007

Lee, K.M. Cho, H.M. & Cha, C.J., Life-cycle cost-effective optimum design of steel bridges considering environmental stressors, *Engineering Structure* 28(9): 1252–1265, 2006

Leeming, M.B., Mouchel, L.G., et al., The application of life cycle costing to bridges, Bridge management 2. Thomas Telford, London, 574–583, 1993

Lim, J.K., Park, H.M. Park, K.H., Cho, H.N. and Kong, J.S., A Practical Life-Cycle Cost Analysis System for Steel Bridge Design and Maintenance Management Steel Bridges, *4th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, FL, USA., 2005

Mark, A. Ehlen., Life-Cycle Costs of New Construction Materials. *Journal of Infrastructure systems*. 3(4): 129–133, 1997

Melchers, R.E., Structural Reliability, Analysis and Prediction, Ellis Horwood Ltd., England, 1987

Office of Construction and Facilities Management, *Data for design guide to maintenance and repair of infrastructure. Final report*, 2002

Piringer, S., Whole-Life Costing of Steel Bridges, Bridge Management, Thomas Telford, London, U.K., 584–593, 1993

Romano, J.J., Modeling Bridge Deck Maintenance Costs, unpublished Master's thesis, Dept. of civil and Environmental Engineering, Lehigh University, Bethlehem, PA., 1997

Stewart M.G. and Hossain, M.B., Time-dependant Deflection, Serviceability Reliability and Expected Cost for RC beams, Structural Safety and Reliability, Corotis et al. (eds), 2001

Walls, J. & Smith, M.R., *Life-Cycle Cost Analysis In Pavement Design – Interim Technical Bulletin*. Fhwa-Sa-98-079. Research Report. Washington, DC: Federal Highway Administration, 1998

Wen Y.K. & Kang Y.J., Minimum lifecycle cost structural design against natural hazards, Structural Research Series No. 629, University of Illinois at Urbana-Champaign, 2000

Yi, S.C., Cho, H.N., Hwang, Y.K. & Lee, K.M. 2003. Practical Life-Cycle Cost Effective Optimum Design of Steel Bridges. *3th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*. Lausanne, Switzerland, 2003

Zimmerman, K.A., Guidelines for Using Economic Factors and Maintenance Costs in Life-Cycle Cost Analysis, Technical Report No. SD96-08-F (NTIS PB98-110638), Applied Pavement Technology, Inc, 1997

Chapter 16

Life-cycle cost and performance prediction: Role of structural health monitoring

Dan M. Frangopol

Department of Civil and Environmental Engineering, ATLSS Research Center, Lehigh University, Bethlehem, PA, USA

Thomas B. Messervey

Department of Structural Mechanics, University of Pavia, Pavia, Italy

ABSTRACT: Monitoring technologies offer great potential to address the aging infrastructure problem and will likely bring with them a new wave of innovations across the design, construction, and management of civil infrastructure. To best leverage this potential, several considerations and actions must be taken that are in contrast to how monitoring is currently most often employed. Instead of a bottom-up reaction to specific deficiency, a top-down approach to the development of monitoring systems within a life-cycle context is necessary. Such an approach requires the adoption of methods and metrics suited for probabilistic data and capable of quantifying the benefit of increased levels of safety over time. It must also be considered that the design, management, and use of civil infrastructure involves a unique composition of interested parties that may compete for resources or have conflicting interests although they share the same goal of safe and efficient structures. To ensure the best use of limited resources, common metrics, methodologies, and means of communication must be agreed upon. Despite the pressing need for new innovations, the integration of structural health monitoring will likely be incremental. As such, how these technologies can benefit existing methods while serving as a catalyst for future change is of interest.

1 Introduction

Structural health monitoring (SHM) is likely the enabling technology that will lead to the next significant evolution of the design, assessment, and management of civil infrastructure. Similar to the impact brought about by computers and structural analysis programs, access to site-specific data across a variety of measurements provides the capability to implement several concepts, methods, and ideas that have existed for some time, but have not yet matured in practical applications. These include, among others, the smart system concept, multifunctional materials, performance and durability based design, life-cycle design, reliability-based structural assessment, and damage detection capabilities. Although effort is typically given to a very specific part of the problem, such as perhaps the design of a particular sensor, a very interesting perspective is brought about by considering how such capabilities will ripple through public policy, code specifications, inspection programs, and educational courses as well as the processes of design and assessment.

Change is a natural and inherent part of civil engineering. The past several decades have witnessed design methodologies shift from deterministic-based approaches, such as allowable stress design, to the semi-probabilistic approaches found in current codes such as American Association of State Highway and Transportation Officials (AASHTO, 2007), the Canadian Highway Bridge Design Code (CHBDC, 2006), and the European Highway Agency Eurocodes (EUROCODES, 2002). In the near future, performance-based design will likely be generally adopted as progress in materials, design software, construction methods, and structural health monitoring empower the structural engineer to better address the uncertainties inherent to the design and operation of civil structures. Although design methods have continued to evolve over time, their intent has remained constant. Each approach (deterministic, semi-probabilistic, and probabilistic) seeks an optimal balance between economical design and safe performance.

Change is typically brought about by opportunity, necessity, or tragedy. At this particular point in time, all three have captured national attention to some extent. Opportunity has presented itself through technological advancements. Reductions in size, wireless capabilities, improved energy performance, and reductions in cost are making SHM practical for civil structure applications. Although monitoring devices have existed for some time, they have typically required a controlled environment, hard wired cables, and immense effort to obtain data making their application in a field environment difficult. Recent improvements in these devices are now making it feasible to obtain site-specific response data cost effectively and offer great potential with respect to the design, assessment, maintenance, and rehabilitation of civil infrastructure (Frangopol and Messervey, 2009). Necessity for the application of such new technologies is most acute for a particular class of civil infrastructure, highway bridges. These structures, in particular, are vulnerable to and are constantly subjected to aggressive environments which include chemical attack from de-icing salts, environmental stressors such as wind, temperature, and water, as well as continuously increasing traffic volumes and heavier truck loads (Frangopol and Liu, 2007). The deterioration of highway bridges in North America, Europe, and Japan is well documented and publicized. In the United States, 25.8% of the 596,808 existing bridges were structurally deficient or functionally obsolete as of the end of 2006 (FHWA, 2007). As the bridge population continues to age and increase, these percentages and numbers are expected to rise unless significant intervention is undertaken. Additionally, because the majority of U.S. bridges were constructed during the 1950–1980s, the U.S. is approaching what some describe as the age of "mass maintenance" with similar trends being reported in both Europe and Asia (Peil, 2003, Fujino and Abe, 2004). To resolve this problem in the United States, an estimated annual investment of 9.4 billion dollars for the next 20 years is necessary (ASCE, 2005). Because so many structures require maintenance, repair, or replacement simultaneously, there is a great need for methods and technologies that accurately assess these structures, prioritize repairs, and enable the efficient allocation of funds. Lastly and unfortunately, the recent collapse of the I35W Bridge in Minneapolis, Minnesota, USA, has focused attention and scrutiny on all aspects of bridge safety to include design practices, considerations for non-redundant structures, inspection and management programs, and the proper conduct of maintenance and repair activities. Although the investigation is still ongoing, the most important lesson from this tragedy may be a reminder of the magnitude and

severity of the indirect costs which include the loss of public confidence and trust, site cleanup expenses, user delays, longer travel distances, decreases in productivity, and legal ramifications.

Developing and leveraging the use of monitoring technologies for civil applications requires insight, planning, and continued research. The rapid pace of advances in monitoring technologies provides a sharp contrast when compared to the time required to affect changes in civil engineering, a field governed by laws, codes, time-tested experience, and where projects themselves may span decades. Although SHM offers great potential, it should be anticipated that such technologies will not be adopted unless they are proven cost-effective due to competing resource demands from the backlog of required maintenance and rehabilitation activities. As such, metrics and methods that calculate and communicate the costs and benefits associated with monitoring must be identified and employed so that alternatives may be adequately compared. With special emphasis on life-cycle cost and performance prediction, this chapter examines these issues and identifies what metrics, methods, and actions are most appropriate for the inclusion of SHM into the design, assessment, and management of civil infrastructure.

2 Life-cycle analysis and durability-based design

Traditional structural design has focused on obtaining the least cost solution that fulfills specified requirements. Cost reduction has typically been gained through reduction in structural weight. Although repairs and upkeep are implicit to all structures, the intended service life of a structure is often left unspecified. In such cases, project bids consider only the initial costs of design and construction and upon completion the structure is turned over to the owner with the absence of a maintenance plan. Maintenance and repair activities are then likely to become an ad-hoc reaction whenever a defect manifests itself (Bijen, 2003). Unfortunately, in terms of expense, research in the field of life-cycle management has shown that the costs of inspecting, maintaining, and repairing a structure over its useful lifespan often dwarf those associated with the initial design. This is compounded by the frequent desire to extend the service life of a structure beyond that originally intended (Estes and Frangopol, 2005).

Aside from routine and anticipated loads, structures may also be subjected to natural hazards such as earthquakes, floods, or hurricanes, as well as other man-made disasters such as vehicle collisions, fires, explosions, or terrorist attacks. Over time, structural safety and condition gradually deteriorate due to normal wear and tear as well as exposure to aggressive environmental stressors. Steel corrodes, concrete spalls, wood rots, and almost all materials crack as they age and progress through their service life. Due to the uncertainties surrounding in-service loads and deterioration processes, the overall performance of the structure may not follow initial predictions as depicted in Fig. 1.

Although the profile may vary, the trend over time reflects the natural aging process where the capacity of the structure to carry loads is reduced and the functionality of the structure may be impaired. Although users typically have a moderate threshold for a decrement in functionality, the failure of any such structure is deemed unacceptable.

In response to these concerns, maintenance and risk mitigation are required to ensure satisfactory performance over the life of a structure as shown in Fig. 2. Because

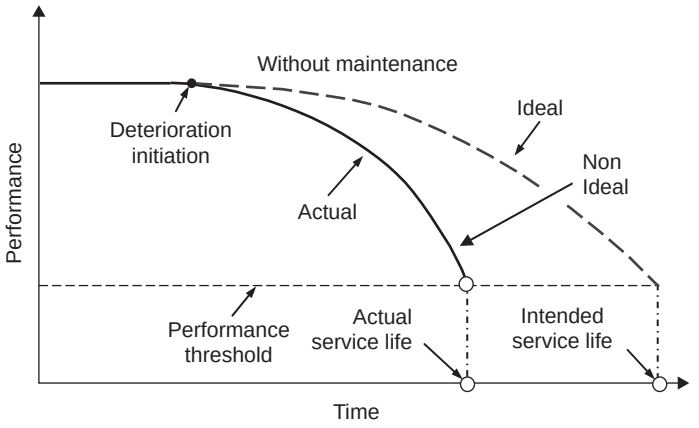


Figure 1 Lifetime structural performance without maintenance (adapted from Frangopol and Liu, 2006).

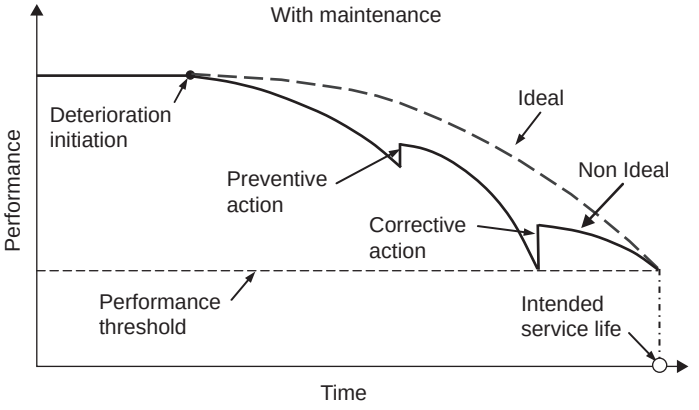


Figure 2 Lifetime structural performance with maintenance (adapted from Frangopol and Liu, 2006).

maintenance needs are often greater than available funds, decisions and scenarios for maintaining infrastructural systems must be based on a life-cycle cost (LCC) analysis (Frangopol, 1998, Frangopol et al., 2001, Kong and Frangopol, 2004). The goal of any such analysis is to cost-effectively allocate resources such that condition, safety, and performance are optimized for individual structures as well as the network within budgetary constraints. This requires the calculation of the present cost of future maintenance and repair actions. Naturally, this implies a prediction of how the structure will behave over time as well as what actions will be taken to mitigate structural deficiencies.

Because structures may be in service for decades or even centuries, there is a great deal of uncertainty with their anticipated loads, deterioration, or even usage. As a result, there is a great need to update life-cycle models with site specific data using inspections,

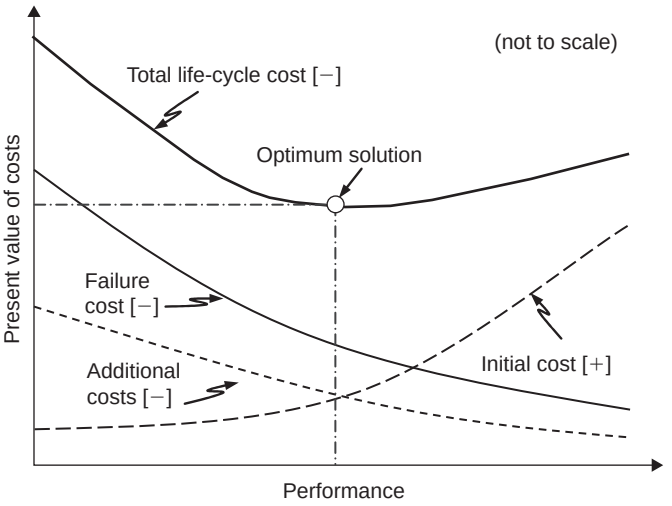


Figure 3 Optimum solution based on life-cycle cost minimization; the sign between the brackets indicates the effect of monitoring on cost (+ and − indicate increase and decrease, respectively) (adapted from Frangopol and Liu, 2007).

non-destructive evaluation, and structural health monitoring. Such updates separate the ideal and actual performance curves depicted in Fig. 1 and Fig. 2.

Against this backdrop, monitoring technologies have the potential to improve the life-cycle management of civil infrastructure in several ways: (a) inspections can be scheduled on an "as needed" basis driven by structure specific data when indicated by monitoring data, (b) the accuracy of structural assessments can be improved by analyzing recorded structural response data, (c) as a result of more accurate information provided as input to analytical models, maintenance, repair, and replacement activities can be optimally scheduled which results in cost savings, and (d) performance thresholds can be established to provide warning when prescribed limits are violated. However, these benefits also come with an associated life-cycle cost as monitoring systems must be purchased, installed, maintained, and their information processed and assessed. Hence, a truly optimal and efficient design needs to consider and evaluate the costs and benefits of different strategies and approaches (Frangopol and Messervey, 2008). Conceptually, the tradeoff between types of costs and their impact on the total life-cycle cost is shown in Fig. 3 with the effect of the inclusion of monitoring on cost denoted in brackets, where [+] and [−] denote additional and less cost respectively. At the extremes of this figure, the lowest possible initial cost (minimum initial specified design) is paired with higher maintenance, repair and inspection costs as well as a higher failure cost. Conversely, the highest initial cost (overdesign) is paired with lower maintenance, repair, and inspection costs as well as lower failure costs. The optimal solution is the one that finds the lowest total life-cycle cost. Although it is possible to design a monitoring system otherwise, the intended goal is to reduce total life-cycle costs through lower additional and failure costs at the expense of higher initial costs.

3 Inclusion of monitoring in life-cycle cost calculations

The general form of the expected life-cycle cost can be calculated as (Frangopol et al., 1997)

$$C_{ET} = C_T + C_{PM} + C_{INS} + C_{REP} + C_F \tag{1}$$

where C_{ET} = expected total cost, C_T = initial design/construction cost, C_{PM} = expected cost of routine maintenance, C_{INS} = expect cost of performing inspections, C_{REP} = expected cost of repairs and C_F = expected cost of failure. Inclusion of monitoring into this general form results in

$$C^0_{ET} = C^0_T + C^0_{PM} + C^0_{INS} + C^0_{REP} + C^0_F + C_{MON} \tag{2}$$

where C_{MON} = expected cost of monitoring which is best treated with respect to a life-cycle cost as

$$C_{MON} = M_T + M_{OP} + M_{INS} + M_{REP} \tag{3}$$

where M_T = expected initial design/construction cost of the monitoring system, M_{OP} = expected operational cost of the monitoring system, M_{INS} = expected inspection cost of the monitoring system, and M_{REP} = expected repair cost of the monitoring system. The benefit of the monitoring system, B_{MON}, can then be determined through a comparison of the expected life-cycle total cost with and without monitoring by subtracting Equation (2) from Equation (1)

$$B_{MON} = C_{ET} - C^0_{ET} \tag{4}$$

Using cost as the sole criterion, monitoring would only be justified if $B_{MON} > 0$ meaning that monitoring is found to be cost effective (Frangopol and Messervey, 2007).

Critical to an appropriate calculation of monitoring utility is the incorporation of risk. One of the main advantages of using monitoring is an increased level of safety through the reduction of uncertainty which must be quantified to compare monitoring vs. non monitoring alternatives. Risk, or the expected cost of failure C_F, can be calculated as the product of the likelihood of an event and the associated consequences given the event occurs as

$$Risk = R = C_F = p_f C \tag{5}$$

but is limited to a point-in-time analysis. Introducing risk into a time-dependent reliability analysis requires the use of a hazard function $H(t)$ to calculate the instantaneous probability of failure or failure rate. This function expresses the conditional probability of failure in time $(t, t + dt)$, given that failure has not already occurred as

$$H(t) = -\frac{dp_s(t)}{dt} \times \frac{1}{p_s(t)} = -\frac{S'(t)}{S(t)} \tag{6}$$

where $p_s(t)$ is the probability that the structure is safe at any time t which is also referred to as the survivor function $S(t)$, and $S'(t)$ its derivative. This function can

then be multiplied by the consequence C (e.g., in US dollars) at each time increment, converted to a net present value, and summed across the lifespan of the structure. Examples can be found in Frangopol and Messervey (2007) and Estes and Frangopol (2005).

4 The development of top-down monitoring strategies

Most often, monitoring is used as a bottom-up, diagnostic tool in response to an existing problem or defect or to conduct system identification for a finite element model. Equipment is brought to the sight, measurements are recorded, the equipment is removed, and the data is studied. In time, as technologies, metrics and methods are developed that are convincingly cost-effective, the use of permanent (or systematic) monitoring systems will become more common. To ensure these assets are employed effectively, they need to be applied at the most critical structure, at the appropriate location, and at the right time.

4.1 *Program level considerations*

Efficient monitoring strategies must begin at the program level. The development of a monitoring program across a fleet of bridges and within an environment of competing demands should consider:

- Funding, Ownership, and Responsibility
- Past Failures and Current Condition
- Importance within a Network

4.1.1 *Funding, ownership, and responsibility*

Bridge ownership can vary across levels of government and political structures. In the United States, these divisions are generally federal, state, county, and local. In contrast, in several European countries, highway bridges are owned and maintained by a private for profit organization funded through tolls. Other bridges fall under their respective municipalities. Such divisions in ownership create the potential for differences in standards, conduct, and clarity with respect to bridge inspections and maintenance. Following the I35W Minneapolis Bridge collapse, the internet news portal MSNBC initiated a series of special reports on the state of the nation's bridges (Dedman, 2008a). One of these reports highlights how the issue can become complex. In this report, several states described how local bridges were not their concern because the structures were owned by the local municipalities. However, such a stance conflicts with federal regulations which state that although the task of bridge inspections can be delegated, the responsibility for their completion cannot be delegated. The report goes on to highlight that although federal officials are aware of many practices that violate federal regulations, no penalty has been levied on any particular state in over 15 years. Restrictions on funding can also complicate the matter. Although experts attest that well allocated maintenance funding significantly reduce life-cycle cost, states and consequently local governments in the United States were until recently precluded from using federal funds for bridge maintenance purposes. Instead, highway gas tax funds

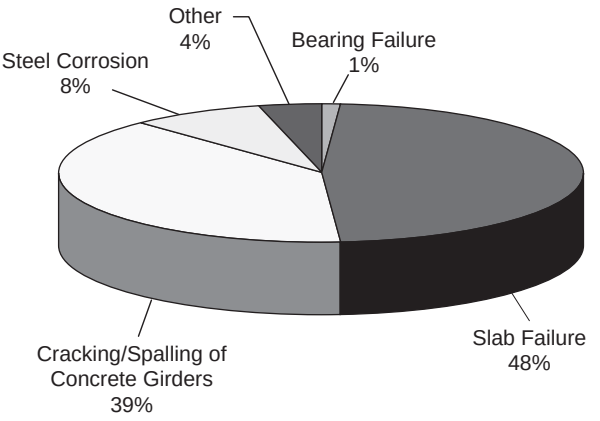


Figure 4 The primary causes of failure for bridge superstructures identified in a 2004 study in Japan (adapted from Fujino and Abe, 2004).

were restricted for new bridge construction and the replacement or rehabilitation of existing bridges (Roberts and Shepard, 2002). As a result, a significant number of local governments and several state governments are only within the last several years beginning to implement bridge maintenance programs. Although it is not the role of a monitoring program to address these issues, it is important to understand the political and funding structure the program must fit within. Additionally, agreements on how to classify monitoring (as part of maintenance or rehabilitation) may directly impact what types of funds are available.

4.1.2 *Consideration of past failures and the condition of existing structures*

Historically, albeit unfortunately, failures and collapses have acted as the catalysts that have shaped design codes, construction methods, and management practices. Several notable studies related to the field of civil engineering are collected in Stewart and Melchers (1997).

Recent and particular to highway bridges is a 2004 review of the reasons for reconstruction across 1691 bridges in Japan (Fujino and Abe, 2004). The results of a part of this study that identify the reasons attributed to bridges with superstructure failures are shown in Fig. 4. From this study, it might be concluded that monitoring strategies for concrete may be of particular interest for bridge managers as slab failure and concrete spalling/cracking accounted for most superstructure failures.

Although past failures provide insight, the current condition and classification of existing bridges must also be considered. In addition, ongoing rehabilitation efforts and construction trends can help shape strategic decisions for the monitoring priorities of new and existing bridge structures. For example, such data may show that one particular type of bridge or construction material is problematic. Although this may justify monitoring priorities for existing structures, this may not carry over to priorities for new structures unless the same type of structure and same type of material is being

used. In the United States, such statistics are compiled in the National Bridge Inventory (FHWA, 2007), which now contains over 40 years of construction records and inspection data. In Frangopol and Messervey (2008), data from the NBI as of 2006 was analyzed to investigate deficient bridges by material type and separately to investigate what materials were predominant in new construction. From this analysis, it was concluded that steel bridges warrant priority in the development of monitoring solutions for existing structures (largest percent deficient) and concrete construction should be targeted for development of monitoring systems for newly constructed bridges (most common type of construction).

4.1.3 *Importance within a network*

To optimize available resources, monitoring should be prioritized to the structure that maximizes utility for the network under consideration. For a transportation network where bridges serve as critical nodes, analysis requires consideration of network connectivity, user satisfaction, and network reliability (Liu and Frangopol, 2006a). Monitoring could be allocated to the most important bridge within a network with respect to any of these three metrics or to a bridge with known defects. An appropriate starting point is to relate individual bridge reliability to the reliability of the bridge network. The reliability importance factor (RIF) for any bridge is defined as the sensitivity of the bridge network reliability β_{net} to the change in the individual bridge system reliability $\beta_{sys,i}$ as (Leemis, 1995, Liu and Frangopol, 2006b):

$$RIF_i = \frac{\partial \beta_{net}}{\partial \beta_{sys,i}} \tag{7}$$

Using this metric, the bridge for which changes in performance have the largest impact on the reliability of the bridge network can be identified for monitoring priority. Assuming that decision based on monitoring reduces the probability of failure of any associated bridge component or system and likewise increases the network reliability index, a multi-objective approach can be utilized to optimize bridge network maintenance as presented in Liu and Frangopol (2006a).

4.2 *Structure level considerations*

At the structure level, monitoring should be allocated to the most important members, for the critical performance functions, to characterize the most significant random variables, at the appropriate point in time. In pursuit of this goal, the development of a structure-level monitoring program should consider the following:

- Types of Monitoring and Measurement
- Role and Treatment of Uncertainty
- Component vs. System Analysis over Time
- Time and Spatial Effects

4.2.1 *Types of monitoring and measurement*

Monitoring strategies are broadly categorized in two groups, global and local. Both provide different types of information and in general support different analysis types.

Selecting an appropriate strategy might be dictated by the structure, type of analysis, or both. Global monitoring may be required when certain parts of a structure are inaccessible or preferred when working with equivalent models such as determining stiffness for a finite element (FE) analysis. Conversely, local monitoring may be preferred when targeting specific random variable parameters, such as live load monitoring through weigh in motion (WIM) data, identified defects, such as crack growth, or specific indicators, such as the presence of chloride ions in the concrete above reinforcing steel. Local monitoring requires close attention to sensor spacing intervals to provide a sufficient probability of detection. Measurements that support these monitoring strategies are typically dynamic or static in nature with global monitoring typically being associated with dynamic structural characteristics and static measurements more often being associated to local monitoring. A more detailed discussion of monitoring strategies and the sensors that support them can be found in (Rafiq, 2005).

4.2.2 *Role and treatment of uncertainty*

Consideration of the uncertainty associated with critical loading and structural parameters is one of the most important issues in assessing the condition of existing civil infrastructures (Catbas et. al, 2008). Uncertainty can be considered in two broad categories, aleatory and epistemic. Aleatory uncertainty describes the inherent randomness of phenomenon being observed and cannot be reduced whereas epistemic uncertainty describes the error associated with imperfect models of reality due to insufficient or inaccurate knowledge (Ang and Tang, 2007). The goal of a monitoring program is to minimize epistemic uncertainty and thus classify the aleatory uncertainty as best possible for treatment within a probabilistic analysis. Both types of uncertainty play an important role in the monitoring of civil infrastructure. System identification, i.e. validating structural parameters through experimental testing, proof loading, or measurements from the structure of interest all act to reduce epistemic uncertainty by improving the accuracy of model input parameters. Efforts should naturally be focused on the parameters for which better information has the greatest impact on model improvement.

An example of recording an aleatory phenomenon (temperature) to reduce the epistemic uncertainty in a system identification model for use in a reliability analysis can be found in Catbas et al. (2007, 2008). Instead, methods for the identification and treatment of uncertainties in the characterization of monitoring-based distributions for use in reliability analyses can be found in Ang (2007), and Messervey and Frangopol (2008a).

4.2.3 *Component vs. system analysis over time*

For structural components (such as bridge decks) and individual members (such as girders), how these elements perform within the context of the larger structure will determine their importance with respect to monitoring. If arranged in series, where the failure of any one element leads to the failure of the system, then the *weakest* element (i.e., the element with the highest probability of failure) is the most important. Conversely, if elements are arranged in parallel, such that the failure of all elements will lead to the failure of the system, then the *strongest* element (i.e., the element with the lowest probability of failure) is the most important. Due to potentially differing

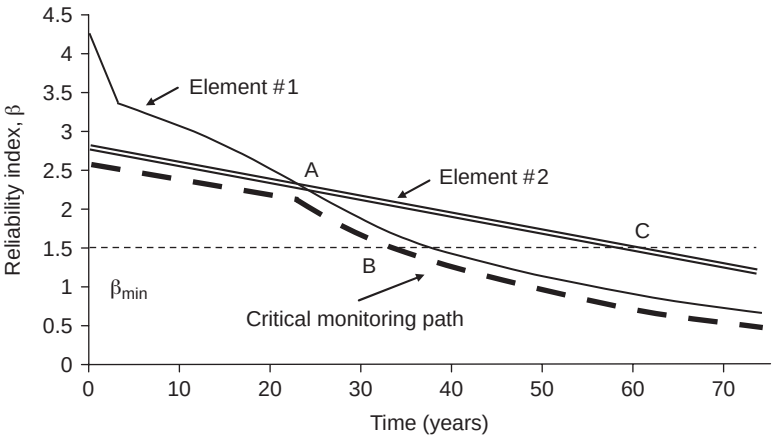


Figure 5 Time Variant System Monitoring of Elements in Series.

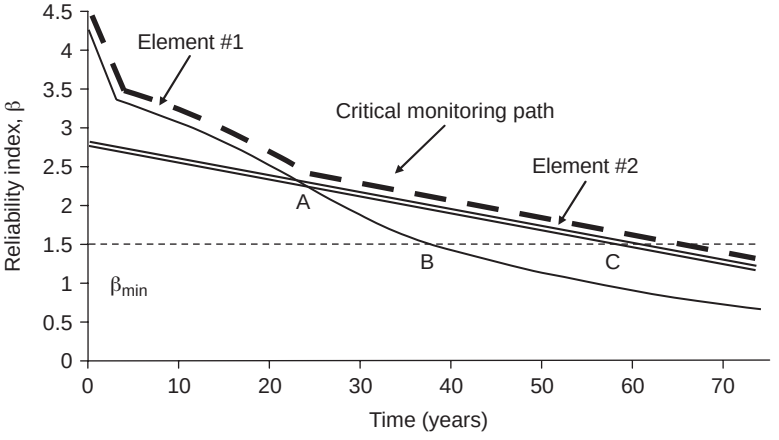


Figure 6 Time Variant System Monitoring of Elements in Parallel.

rates of deterioration amongst members, as is often the case for exterior members more exposed to environmental stressors that lead to corrosion, monitoring priorities may shift between members or components over time as the structure deteriorates.

These concepts are illustrated by examining the reliability profiles of two members considered first in series, then in parallel. Considered in series as shown in Fig. 5, Element #2 is most important because of its lower reliability index until Point A is reached at which time Element #1 takes priority due to a faster rate of deterioration.

Considered in parallel as shown in Fig. 6, the opposite scenario is present where Element #1 is at first more important due to its higher reliability index until Point A is reached and Element #2 takes priority. Hence, depending upon the age of the structure, member configuration, and the type of monitoring system employed (permanent vs. non permanent), monitoring priority may be appropriate for either member.

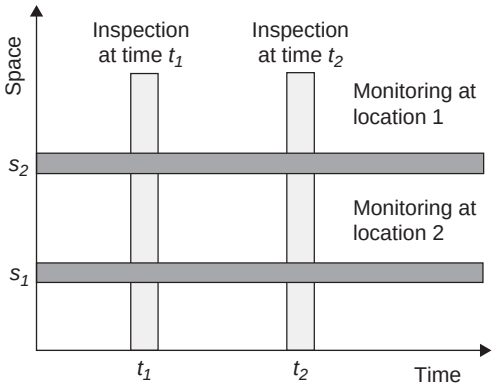


Figure 7 Temporal and spatial differences between inspections and monitoring (adapted from Rafiq, 2005).

Often, maintenance and repair actions are driven by a minimum reliability threshold, β_{min}. Once this value is established, a time dependent reliability analysis can also answer the question of when to monitor. Using $\beta_{min} = 1.5$ as shown in both Fig. 5 and Fig. 6, monitoring would be appropriate on Element #1 at the time corresponding to Point B for elements in series (Fig. 5) and on Element #2 at Point C for elements in parallel (Fig. 6). In the absence of perfect information, Monte Carlo simulation can be utilized to estimate the earliest possible crossing of the minimum reliability threshold (Neves et al., 2006). This method is appropriate for a monitoring system with high operational costs that can be turned on or off, or for a non-permanent monitoring solution that must be scheduled and brought to the site.

4.2.4 *Time and spatial effects*

Fig. 7 depicts an important difference between inspections and monitoring. Inspections classify the state of an entire structure at a point in time whereas continuous monitoring provides insight as to a particular location throughout time. Employing both inspections and monitoring for the maintenance of a bridge structure, the goal is to cost-effectively achieve adequate coverage both temporally and spatially. To this end, Marsh and Frangopol (2007) employ multi-objective optimization to design a sensor network in a reinforced concrete deck based upon cost, sensor performance, and measurement uncertainty. In general, the idea is to record enough information to extrapolate to other parts of the structure in a manner that is statistically significant while balancing increased information vs. its associated cost.

Time effects must also be accounted for in order to relate observed monitoring data to code specified levels of safety thus distinguishing serviceability from ultimate limit states. For example, if a structure is monitored for a short period and only light load demands are recorded, one cannot conclude that the structure has enough safety. One could conclude that the structure *was safe* for the loads encountered during the period monitored, but does not adequately convey the safety of the structure for its intended

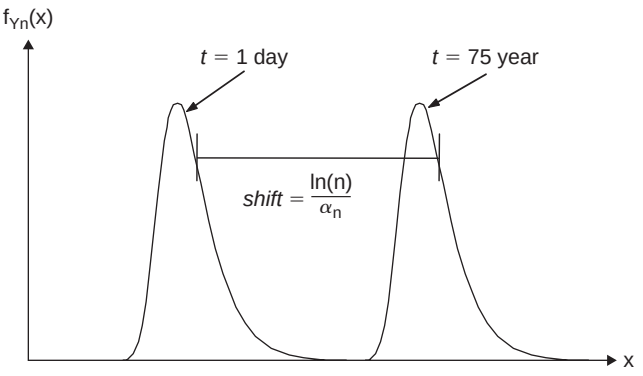


Figure 8 Transformation of a Type I Gumbel Distribution from a daily observation timeframe to a 75-year timeframe.

service life. When considering the entire life of a structure, it becomes clear that extreme events such as combinations of overloaded trucks, hurricanes, earthquakes, and other foreseen events must be considered. Specific to the development of a monitoring-based live load for use in a reliability analysis, Messervey and Frangopol (2008b) demonstrate how the asymptotic behavior of extreme value distributions can relate information observed in one timeframe to another timeframe of interest. In this work, a Type I Gumbel distribution is characterized by peak picking maximum values and is then transformed to the 75 year return period for live loads consistent with the period used for LRFD code calibration (Ghosn et al., 2003). Conceptually, this transformation is illustrated in Fig. 8.

5 Incorporating monitoring into design and management

The design and management of civil infrastructure are unique in that the end product is one of public service instead of profit. In contrast to private industry where new technologies may be adopted very quickly, incorporating new advancements in the field of civil engineering may take considerably longer as laws, codes, and political processes are involved. Within this environment, there is a great need for cooperative, coordinated and synergistic actions by scientists, engineers, infrastructure managers, and governing officials. Technologies must be adopted by and work within the programs utilized for asset management. In turn, asset management must be supported by and exist within the broader context of performance-based engineering. Instrumental and inherent to the entire hierarchy is that resources are optimized, safety is assured, and condition is adequate. Because such an effort will require time, trial and error, and an immense amount of data collection, it is desirable that new innovations fit into existing design methods and management programs for implementation now while serving as a catalyst for the development of new approaches specifically suited for in-service data (Messervey and Frangopol, 2008c).

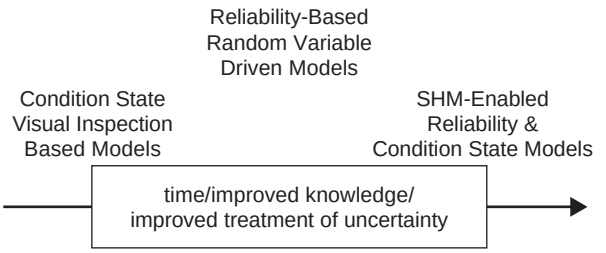


Figure 9 Projected Evolution of Bridge Management Programs.

5.1 *Using monitoring to improve assessment and management programs*

Several bridge failures in the 1960s focused national attention on bridge safety resulting in the initiation and standardization of federally mandated bridge inspections established with the Federal Highway Act of 1968. Since this time, the Federal Highway Administration (FHWA) has revised national bridge inspection standards (NBIS) almost yearly as methods and the base of knowledge in the field have improved. Currently and almost exclusively, bridge management programs are based on visual inspections. In special cases, non-destructive evaluation (NDE) tests are performed to investigate a specific area or problem of interest. Although relatively few, there is a growing number of monitoring applications. Qualitatively, Fig. 9 depicts the anticipated evolution of bridge management programs by type as the capability to obtain data and to treat uncertainty increases. At present, the current state-of-the art is toward the left side of the figure reflecting the predominant use of visual inspection based condition state models.

Condition state models such as Pontis (Pontis, 2007) are currently the most widely adopted bridge management programs in service. Based primarily on visual inspection data, the main advantage of these models is that they are relatively simple to implement. The primary limitation of condition-state models is that safety is not adequately addressed as visual appearance does not always correlate to structural performance and accuracy is lost due to a limited number of discrete condition states (Frangopol and Liu, 2007). Human error is also a consideration. One recent study reported that in some cases more than 50% of bridges are being classified incorrectly via visual inspections (Catbas et al., 2007). Although certainly not the norm, a separate recent article highlighted the falsification of bridge inspections by contractors to keep up with timelines for reporting purposes (Dedman, 2008b). In response to these limitations and concerns, reliability-based models were developed that specifically assess structural safety (Watanabe et al. 2004, Casas et al. 2002). Such models seek to incorporate all random variables and parameters affecting structural performance and are well suited to account for structural redundancy through system analysis. However, due to the amount of input parameters required and the sensitivity of the results to the accuracy of the input, these models have been limited in implementation. In the foreseeable future, it is likely that monitoring technologies will provide the mechanism to make reliability-based models more practical for implementation. A further improvement

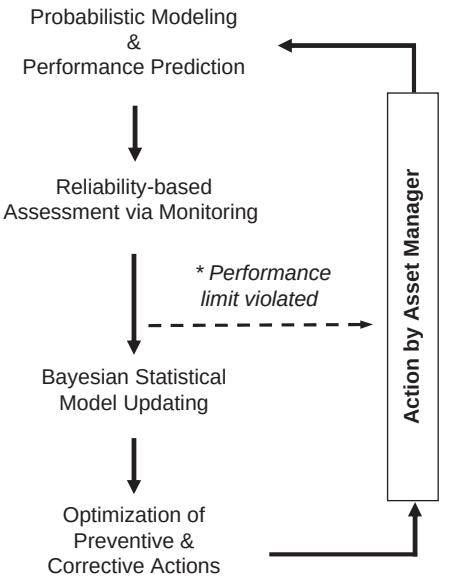


Figure 10 General framework for a combined LCM/SHM approach to infrastructure management.

on reliability-based models is possible by combining the advantages of the condition-state and reliability-based models. As condition-state models do not directly address safety, reliability-based models do not directly address condition where repairs may be required to improve trafficability despite a high level of structural safety. This has led to the development of hybrid-type models that account for both condition and safety (Frangopol, 2003, Neves et al. 2006, Bucher and Frangopol, 2006, Neves and Frangopol, 2004). Such models provide a more holistic treatment of the problem but also imply a greater degree of complexity and increased cost.

Although monitoring can benefit any of these bridge management models, it is best suited for the reliability-based or hybrid types. Ideally, a bridge management program allows for individual or bridge network assessment, maintenance, inspection, and repair planning based on real-time structure-specific data. Such approaches are heavily dependent upon monitoring data. How to best integrate this data into reliability-based life-cycle management programs is being investigated by many researchers worldwide (Messervey and Frangopol, 2007, Budelmann and Hariri, 2006, Klinzmann et al., 2006). A combined approach to bridge management that incorporates LCM and SHM is attractive as the advantages of each approach offset the other's disadvantages. LCM offers bridge managers a practical predictive view of cost, safety, and condition, but in many regards lacks knowledge of actual structural performance. In contrast, SHM effectively captures structural behavior and load demands, but is not as effective in translating this information into actionable information. Fig. 10 depicts a general framework for the inclusion of SHM in LCM models. The process begins with a reliability-based treatment of the structure that helps determine a task-oriented monitoring solution and initial performance prediction. Once in place, an assessment loop

begins in which monitoring data is utilized to update the structural model. Once the predictive model is updated, maintenance actions can be optimized for decision by the asset manager. If the monitoring solution employed is permanent (continuous), performance flags can serve as a warning system to alert the asset manager immediately of any violation of a critical threshold or unexpected distress. The result of any such integrated LCM/SHM approach is an adaptive, self-learning management system with the capacity to improve the underlying theoretical-based models through structure specific response data over time. A more accurate model provides the potential for cost savings through optimal maintenance and inspection scheduling, and a decrease in risk through the reduction of uncertainty.

5.2 *Using monitoring to improve design methods*

Although health monitoring is associated almost exclusively with the assessment of in-service structures, SHM also has great potential to improve design codes and methods. As a natural an inherent part of the design process, the assessment of existing structures, failures, and experiments has historically provided the knowledge base for the improvement of existing techniques. As such, monitoring has great potential to improve design if the information is collected and leveraged for future code revisions. This is not a new idea, but rather a natural extension of what has already been occurring to ensure that the codes are living documents that serve the best interest of society.

Similar to bridge management programs, design methods have evolved over time as the ability to model and account for uncertainty has improved as depicted in Fig. 11.

Although some improvement is possible for allowable stress approach, the inclusion of monitoring data is best suited for both the factored and performance based approaches as monitoring data is uncertain in nature. For factor-based semi-probabilistic approaches (LRFD), monitoring can provide data to confirm or improve existing load factors, resistance factors, and load combinations for extreme events. Many notable and complex studies have been undertaken to model the performance of in-service bridges over time (Ghosn and Moses, 1998, Enright and Frangopol 1999a, 1999b, 1999c, Ghosn, 2000, Ghosn et al. 2003, Gindy and Nassif, 2006). Such studies have sought to better model truck populations, the distribution and frequency of traffic configurations on structures, load distribution within structures, system effects, and the combinations of extreme events. Monitoring presents the opportunity to revisit such studies, improve modeling assumptions and parameters, and recalibrate factors to achieve more accurate design codes.

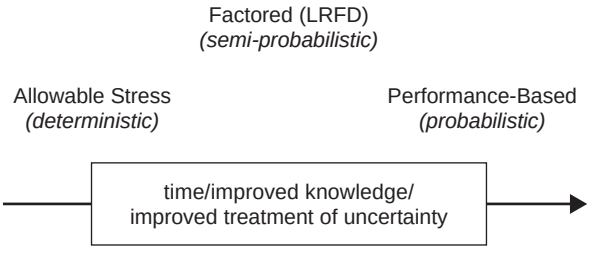


Figure 11 Evolution of Design Methods.

Although there is much room for improvement in existing codes, the true potential of monitoring information is in enabling the adoption and use of performance-based design. Based upon structural reliability, in performance-based approaches the engineer is responsible for the specification of all random variables that determine member resistances, load effects, and load combinations. By providing the statistical information necessary to adequately define these parameters, monitoring may serve as the catalyst for a change in the methodology itself. Although fully probabilistic-based design has been proposed in several countries and some provisions allow more flexibility for its use, to date this method has not been widely adopted.

Collecting the information necessary to re-examine existing codes or to provide consistent guidelines for the application of performance-based design is an interesting challenge. Data is required from multiple structure types operating in different environmental conditions over long periods of time. As a result, engineering experts will need to come to a consensus for data collection methods and the appropriate models for their use. Due to cost and the amount of information required, such a collection effort will likely require coordination from national and international agencies. An example of the successful implementation of one such program is found in the field of structural health monitoring for damage detection. Within this field, a benchmark study of a progressively damaged frame structure has been established and data was posted online to provide researchers with a common reference to develop damage detection algorithms (ASCE, 2000). Although replicating such a program for bridges is of substantial effort, the potential benefit is also significant.

5.3 *Macro-level* adoptions-in-concert *to facilitate the development of SHM systems*

Although the goal of optimally designed and managed structures that ensure public safety over their useful lifespan is most likely not disputed, determining how to achieve this goal is more difficult. Differences in methods, assessment metrics, competing interests, and competing demands amongst the following interested parties

- Public Officials
- Users
- Bridge Managers
- Engineers
- Researchers and Educators

quickly complicate the discussion.

In an environment of limited resources, researchers might request funding to develop more efficient management techniques whereas infrastructure managers would prefer using this funding to repair existing defects. The trucking industry desires heavier allowable truck weights to improve productivity whereas bridge managers desire lower limits to reduce wear and tear on their structures. Public officials are responsible for public safety but must also accept some level of risk as its elimination is not feasible or affordable.

Because monitoring technologies enter this environment of limited resources amongst interested parties with potentially conflicting objectives, coordinated and synchronized actions (*adoptions-in-concert*) are necessary to facilitate synergistic and

efficient solutions. Particular attention is needed to standardize and include risk in the calculation of life-cycle costs and as a metric to compare alternatives that do and do not employ SHM. Also necessary is a common period of time (warranty period) over which to calculate maintenance costs for newly constructed structures. In doing so, solutions with and without monitoring could be fairly compared. An action that would synchronize and gain efficiencies between interested parties is the establishment of a focused data collection effort. Although much information is required to update codes and to better understand in-service structures, it is likely that specific failure modes across certain types of bridges are of greatest concern at this point in time. Focusing research funding and effort on identified problem areas and making the data available to different researchers for benchmarking would result in more coordinated, efficient, and timely solutions.

6 Conclusions

Although the challenges associated with the management and repair of aging and deteriorating structures are substantial, technological advances in monitoring technologies offer great potential to address the problem. Furthermore, if leveraged appropriately, SHM has the potential to become the catalyst that brings changes in design and assessment methodologies. With respect to the development of monitoring programs, a top-down treatment of the problem is appropriate to maximize the benefit of limited available resources. Such programs must be innovative in nature and employ SHM technologies at the right place and at the right time to calculate and communicate commonly accepted metrics with respect to the assessment, performance prediction, and cost estimation for both new and existing structures. Because the safe performance of highway bridges spans differences between interested parties with potentially conflicting interests, coordinated and synchronized actions are necessary to both support and facilitate progress in the development of SHM systems. Such macro-level adoptions-in-concert ensure that research efforts are supported and focused to maximize utility for each party involved.

Acknowledgments

The support from (a) the National Science Foundation through grants CMS-0638728 and CMS-0639428, (b) the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA), (c) the U.S. Federal Highway Administration Cooperative Agreement Award DTFH61-07-H-00040, and (d) the U.S. Office of Naval Research Contract Number N00014-08-1-0188 is gratefully acknowledged. The opinions and conclusions presented in this chapter are those of the authors and do not necessarily reflect the views of the sponsoring organizations.

References

Ang, A.H.-S. and Tang, W.H., *Probability Concepts in Engineering*, 2nd Edition, J. Wiley & Sons, Inc., New York, 2007

Ang, A.H.-S., An application of quantitative risk assessment in infrastructures engineering. *Proceedings CECAR 4*, Taipei, Taiwan, 2007

American Association of State Highway and Transportation Officials (AASHTO). LRFD Bridge Design Specifications, 4th ed., Washington, D.C., 2007

ASCE, Structural Health Monitoring Committee, 2000. Benchmark studies. Available online at http://cive.seas. wustl.edu/wusceel/asce.shm/benchmarks.htm (accessed June 2007)

ASCE, 2005. Report card for America's Infrastructure, American Society of Civil Engineers, Reston, VA. Available online at www.asce.org/reportcard/2005/index.cfm (accessed 22 May 2007)

Bijen, J., *Durability of Engineering Structures: Design, Repair and Maintenance*. Woodhead Publishing Limited, Abington Hall, Cambridge, England, 2003

Bucher, C. and Frangopol, D.M., Optimization of lifetime maintenance strategies for deteriorating structures considering probabilities of violating safety, condition, and cost thresholds. *Probabilistic Engineering Mechanics*, Elsevier, 21(1), 1–8, 2006

Budelmann, H. and Hariri, K. 2006, A Structural Monitoring System for RC-/PC-Structures, *Proceedings LCC5*, Seoul, Korea, 2006

Canadian Highway Bridge Design Code (CHBDC). Ministry of Transportation, Ontario, Canada, 2006

Casas, J.R., Frangopol, D.M., and Nowak, A.S., eds., *Bridge Maintenance, Safety and Management*, CIMNE, Barcelona, Spain, 2002

Catbas, N.F., Zaurin, R., Frangopol, D.M., and Susoy, M., System Reliability-Based Structural Health Monitoring with FEM Simulation. *Proceedings SHMII-3*, Vancouver, B.C., Canada, 2007

Catbas, F.N., Susoy, M., and Frangopol, D.M., Structural Health Monitoring and Reliability Estimation: Long Span Truss Bridge Application with Environmental Monitoring Data, *Engineering Structures*, Elsevier 30(9), 2347–2359, 2008

Dedman, B., Late Bridge Inspections Put Public at Risk, MSNBC Online available at http://www.msnbc.msn.com/id/20998261/ accessed 24 March, 2008a

Dedman, B., Ga. Employee Faked Bridge Inspections, MSNBC Online available at http://www.msnbc.msn.com/id/23020686/ accessed 24 March, 2008b

Enright, M.P. and Frangopol, D.M., Maintenance planning for deteriorating concrete bridges. *Journal of Structural Engineering*, ASCE,125(12), 1407–1414, 1999a

Enright, M.P. and Frangopol, D.M., Condition prediction of deteriorating concrete bridges using Bayesian updating. *Journal of Structural Engineering*, ASCE, 125(10), 1118–1124, 1999b

Enright, M.P. and Frangopol, D.M., Reliability-based condition assessment of deteriorating concrete bridges considering load redistribution, *Structural Safety*, 21(2), 159–195, 1999c

Estes, A.C. and Frangopol, D.M., Life-cycle evaluation and condition assessment of structures. Chapter 36 in *Structural Engineering Handbook, 2nd Edition*, W-F. Chen and E.M. Lui, eds., CRC Press: 36–1 to 36–51, 2005

EUROCODES, Basis of design and actions on structures, BS EN 1990:2002. British Standards Institution, London, United Kingdom, 2002

Federal Highway Administration (FHWA), National Bridge Inventory, available online at http://www. fhwa.dot.gov/bridge/nbi.htm (accessed 22 May 2007)

Frangopol, D.M., A Probabilistic Model Based on Eight Random Variables for Preventive Maintenance of Bridges, Progress meeting on optimum maintenance strategies for different bridge types, The Highways Agency, London, 1998

Frangopol, D.M., Preventative Maintenance Strategies for Bridge Groups – Analysis, Final Project Report to the Highways Agency, London, U.K., March, 2003

Frangopol, D.M., Lin, K-Y., and Estes, A.C., Life-cycle cost design of deteriorating structures, *Journal of Structural Engineering*, ASCE, **123**(10), 1390–1401, 1997

Frangopol, D.M., Kong, J.S., and Gharaibeh, E.S., Reliability based life cycle management of highway bridges, *Journal of Computing in Civil Engineering,* ASCE 15(1), 27–34, 2001

Frangopol, D.M. and Liu, M., Life-cycle cost and performance of civil structures. *McGraw-Hill 2006 Yearbook of Science and Technology*, McGraw-Hill, New York: 183–185, 2006

Frangopol, D.M., Liu, M., Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost, *Structure and Infrastructure Engineering*; 3(1), 29–41, 2007

Frangopol, D.M. and Messervey, T.B., Risk assessment for bridge decision making, *Proceedings of the 4th Civil Engineering Conference in the Asia Region*, Taipei, Taiwan, 2007

Frangopol, D.M. and Messervey, T.B., Effect of SHM on reliability of bridges. Chapter 25 in *Monitoring Technologies for Bridge Management: State-of-the-Art*, A. Mufti and B. Bahkt, eds., Multi-Science Publishing Co., Ltd., England (invited paper, submitted), 2008

Frangopol, D.M. and Messervey, T.B., Maintenance principles for civil structures. Chapter 89 in *Encyclopedia of Structural Health Monitoring*. Ed. Christian Boller, Fu-Kuo Chang, and Yozo Fujino. John Wiley & Sons, Ltd., UK, Vol. 4, 1533–1562, 2009

Fujino, Y. and Abe, M., Structural Health Monitoring – Current Status and Future. *Proceedings of the Second European Workshop on Structural Health Monitoring*, Munich, Germany, 2004

Ghosn, M. and Moses, F., Redundancy in highway bridge superstructures. NCHRP TRB Report 406, Washington, D.C., U.S.A., 1998

Ghosn, M., Development of Truck Weight Regulations Using Bridge Reliability Model, *Journal of Bridge Engineering, 5*(4): 293–303, 2000

Ghosn, M., Moses, F. and Wang, J., Design of Highway Bridges for Extreme Events, NCHRP TRB Report 489, Washington, D.C., 2003

Gindy, M. and Nassif, H., Effect of bridge live load based on 10 years of WIM data. *Proceedings IABMAS'06*, Porto, Portugal, 2006

Klinzmann, C., Schnetgöke, R., and Hosser, D., A Framework for Reliability-Based System Assessment Based on Structural Health Monitoring, Proceedings of the 3rd European Workshop on Structural Health Monitoring, Granada, Spain, 2006

Kong, J.S. and Frangopol, D.M., 2004. Cost-Reliability Interaction in Life-Cycle Cost Optimization of Deteriorating Structures, *Journal of Structural Engineering,* ASCE 130(11): 1704–1712.

Leemis, L.M., *Reliability: Probabilistic Models and Statistical Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1995

Liu, M. and Frangopol, D.M., Optimizing Bridge Network Maintenance Management under Uncertainty with Conflicting Criteria: Life Cycle Maintenance, Failure, and User Costs, *Journal of Structural Engineering,* ASCE, 131(11), 1835–1845, 2006a

Liu, M. and Frangopol, D.M., Probability-Based Bridge Network Performance Evaluation, *Journal of Bridge Engineering*, ASCE, 11(5), 633–641, 2006b

Marsh, P.S. and Frangopol, D.M., Reinforced concrete bridge deck reliability model incorporating temporal and spatial variations of probabilistic corrosion rate sensor data. Reliability Engineering & System Safety, Elsevier, 93(3), 394–409, 2008

Messervey, T.B. and Frangopol, D.M., A framework to incorporate structural health monitoring into reliability-based life-cycle bridge management models. *Proceedings CSMC5*, Rhodos, Greece, June 21–23, 2006; in *Computational Stochastic Mechanics*, G. Deodatis and P.D. Spanos, eds., Millpress, Rotterdam, 463–469, 2007

Messervey, T.B. and Frangopol, D.M., Probabilistic Treatment of Bridge Monitoring data and Associated Errors for Reliability Assessment and Prediction, *Proceedings, IALCCE08*, Varenna, Lake Como, Italy, June 11–14, 2008; in *Life-Cycle Civil Engineering*, F. Biondini and D.M. Frangopol, eds., CRC Press/Balkema, Taylor & Francis Group plc, London, 391–396, 2008a.

Messervey, T.B. and Frangopol, D.M., Innovative Treatment of Monitoring Data for Reliability-Based Assessment, *Proceedings IABMAS'08,* Seoul, Korea, July 13–17, 2008; in *Bridge Maintenance, Safety, Management, Health Monitoring and Informatics*, H-M. Koh and

D.M. Frangopol, eds., CRC Press/Balkema, Taylor & Francis Group plc, London, 319, and full paper on CD-ROM, Taylor & Francis Group plc, London, 1466–1474, 2008b

Messervey, T.B. and Frangopol, D.M., Integration of Health Monitoring in Asset Management in a Life-Cycle Perspective, *Proceedings IABMAS'08,* Seoul, Korea, July 13–17, 2008; in *Bridge Maintenance, Safety, Management, Health Monitoring and Informatics*, H-M. Koh and D.M. Frangopol, eds., CRC Press/Balkema, Taylor & Francis Group plc, London, 391, and full paper on CD-ROM, Taylor & Francis Group plc, London, 1836–1844, 2008c.

Neves, L.C. and Frangopol, D.M., Condition, safety, and cost profiles for deteriorating structures with emphasis on bridges, *Reliability Engineering and System Safety*, Vol. 89, pp. 185–198, 2004

Neves, L.A.C., Frangopol, D.M., and Petcherdchoo, A., Probabilistic lifetime-oriented multi-objective optimization of bridge maintenance: combination of maintenance types. Journal of Structural Engineering, ASCE, 132(11), 1821–1834, 2006

Peil, U., Life-Cycle Prolongation of Civil Engineering Structures via Monitoring, *Proceedings of the 4th International Workshop on Structural Health Monitoring*, Stanford, CA, 2003

PONTIS, user manual and product web site, Cambridge Systematics, available through the AASHTO products office at http://www.camsys.com/tp_inframan_pontis.htm (accessed 2 Nov 2007)

Rafiq, M.I., Health Monitoring in Proactive Reliability Management of Deteriorating Concrete Bridges, University of Surrey, UK., 2005

Roberts, J.E. and Shepard, R., Bridge Management for the 21st Century. *Proceedings, IABMAS'02*, Barcelona, Spain, 2002

Stewart, M.G. and Melchers, R.E., *Probabilistic Risk Assessment of Engineering Systems*, Chapman & Hall, London, 1997

Watanabe, E., Frangopol, D.M., and Utsunomiya, T., eds., *Bridge Maintenance, Safety, Management and Cost*, A.A. Balkema, Swets & Zeitlinger B.V., Lisse, The Netherlands, 2004

Chapter 17

Structural health assessment using noise-contaminated minimum dynamic response information

Achintya Haldar

Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, AZ, USA

ABSTRACT: A novel nondestructive structural health assessment (SHA) technique, known as the generalized iterative least squares – extended Kalman filter – with unknown input (GILS-EKF-UI) method, is being developed at the University of Arizona. The primary objective of the procedure is to detect defects in new, deteriorated or rehabilitated existing structures or just after large natural (strong earthquakes, high wind, etc.) or manmade (blast, explosion, etc.) events. The method is essentially a time domain finite element-based system identification (SI)-based procedure to locate defect at the local element level. Most SI-based SHA approaches use excitation and response information to identify a structure. Excitation information is not available in most cases. Furthermore, outside the control laboratory environment, the collection of excitation information could be so error-prone that the SI concept may not be applicable. It will be desirable if a system can be identified without excitation information. For large complicated real structures, it may not be possible to measure responses at all dynamic degrees of freedom and they always contain noise. Addressing all the issues, a Kalman filter-based algorithm is being developed with considerable success. It can be used for rapid diagnostic purpose as a part of a broader SHA and maintenance strategy.

1 Introduction

Structural health assessment (SHA) has become an important responsibility of the engineering profession after observing dramatic collapse of an about 40 years old steel bridge in Minneapolis, Minnesota on September 1, 2007. Apart from other damages, the failure caused 13 deaths and 145 injuries. In early 2008, the National Transportation Safety Board (NTSB) concluded that the improper gusset plate design at the connections was the main reason for the collapse. The flawed gusset plate design cannot be considered as natural aging process and this type of defect may not be detected during routine inspections. However, the bridge was inspected in June 2006 (Minnesota, 2006). Presence of several defects including surface rust, corrosion, pitting, missing bolts, several types of cracks, sliding of plates, etc. were documented (Haldar, Martinez-Flores, and Katkhuda, 2008). Obviously, past inspections failed to identify the major flaws and no corrective action was taken. There are over 4200 bridges in the U.S. that are in the same or worse condition than this bridge. After the Northridge earthquake of 1994, patients from one damaged hospital were moved to another hospital without assessing the level of damage experienced by the relocated hospital. During the same seismic event, welds in steel connections fractured in more than 200 buildings. Similar connections also fractured during the 1989 Loma Prieta

earthquake; however, they were not detected during the 5 years period between the two earthquakes (Mehrabian and Haldar, 2002).

The structural health also needs to be assessed just after a natural event like high wind or man-made event like explosions or blasts. The challenge is how an engineer objectively assesses health of structures under various conditions or situations. Since all structures are expected to have some defects, simply stating that the structure is defective will not address the problem. Since all defects are not equally critical, their locations and criticality will also need to be established as a part of an overall maintenance strategy.

There is no doubt that the profession has technical sophistication to investigate an already identified defect spot. Obviously, for a large structural system, locating the defect spot may not be simple, although in some cases, the potential locations can be established based on experience. The author and his team have been working on SHA for over three decades. Initially, they assumed that the defect locations were known and proposed a fatigue crack detection procedure using ultrasonic test results (Zhao and Haldar, 1994; Zhao, Haldar, and Breen, 1996). They proposed a technique to update the risk as results from imperfect inspection become available and developed a decision analysis frame work to suggest what to do after an inspection when cracks were detected with sizes measured and unmeasured or not detected during the inspection. For large structural systems, acceptable inspection procedures are not available, the location(s) and types of defect will be unknown in most cases, and the cost of inspections including restrictions imposed on their usage will be so large that it will be economically unacceptable. Bridges are expected to be inspected regularly, say about every two years, most of the time visually. The visual inspections may not be effective if the defects are hidden behind obstructions or inaccessible. Visually detecting one defect does not assure that all defects are detected. The discussions clearly indicate that a new SHA technique is urgently needed for different types of large structural systems.

2 Desirable features of a new SHA

Static or dynamic responses can be used or preferred to objectively assess in-service deteriorated health reflecting the current structural behavior instead of any analytical investigation. In general, dynamic responses are expected to give more complete information on the structural health. Dynamic responses, even measured by smart sensors, are expected to be noise contaminated. Furthermore, for large structural systems, it may not be possible or economical to measure dynamic responses at all dynamic degrees of freedom (DDOFs). Thus, the ideal procedure should be able to detect defects using less than ideal number of responses. If the defects need to be identified at the local element level, it will be highly desirable to represent the structure by finite elements. By tracking the dynamic properties (mass, stiffness, and damping) of each element and comparing them with the expected properties or values suggested in the design drawings or changes if periodic inspections are conducted or simply comparing the properties with each other, the defect spots and their severity can be established.

To satisfy some of these desirable features, an inverse transformation technique known as the system identification (SI) process can be used to assess structural health. The most basic SI approach has three elements: (1) the excitation that caused the

response, (2) the system to be identified, and (3) the measured responses. If the information on the excitation and the response is available, the dynamic properties of the system can be identified. Since the information on mass of elements is readily available, it is assumed to be known in most identification problems (Wang and Haldar, 1994). Thus, the identification process is essentially the estimation of the stiffness and damping properties of all the elements in the finite element representation.

For the sake of completeness, it needs to be stated that currently available SI-based defect identification techniques can be subdivided in to two groups: (i) the frequency domain or modal approaches and (ii) the time domain approaches. In modal approaches, frequencies and mode shape vectors are used to detect defects. An overwhelming majority of researchers use this approach. However, it has been established by the author and others that the changes in the frequencies and mode shape vectors are not sensitive to the defects. A structural member can be completely broken; the changes in the frequencies could be only of 2%. Similar changes can be expected in a defect-free structure due to presence of noises in the responses. Furthermore, modal properties are based on structural global properties. Frequency based SI approaches can evaluate defective state in the overall sense, i.e., whether the structure is defective or not without identifying the defect spots. Although, comparatively more difficult, a time domain approach is preferable. The proposed method is such an approach.

An additional challenge is that measured excitation and responses in time domain are always noise contaminated. Outside the control laboratory environment, the measurement of excitation could be very challenging or so error-prone that the SI concept may not be applicable. Also, after a strong earthquake or high wind, the excitation information for a specific structure may not be available. To increase the application potential of the proposed method, a system needs to be identified without using any excitation information. This is expected to be very challenging since two of the three unknowns in the SI-process will be unknown.

In summary, an intelligent performance-based time domain finite element-based SI procedure that can assess the health of structural elements using only limited noise-contaminated response information is needed as a part of a broader structural health assessment and maintenance strategy. However, Maybeck (1979) commented that a deterministic system cannot be identified using measured response information. The author, with the help from his associates, is in the process of developing such a technique. They called it the Generalized Iterative Least Squares – Extended Kalman Filter – Unknown Input (GILS-EKF-UI) method (Katkhuda and Haldar, 2007). It is the subject of this paper. Analytical and limited experimental verifications of the method are briefly discussed in this paper.

3 Concept behind ILS-UI and MILS-UI approaches

For the easier presentation of the GILS-EKF-UI method, a chronological development of the basic concept is necessary.

3.1 *Iterative least squares with unknown input (ILS-UI) method*

Incorporating some of the desirable features of the novel concept discussed in Section 2, the research team initially developed a method known as the Iterative Least

Squares with Unknown Input (ILS-UI) method (Wang and Haldar, 1994). Without loosing generality, the general governing dynamic equation of motion of a system can be expressed in a matrix form as:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{f}(t) \tag{1}$$

where **M**, **C**, **K** are time-invariant global mass, damping, and stiffness matrices, respectively; $\ddot{\mathbf{u}}(t)$, $\dot{\mathbf{u}}(t)$, $\mathbf{u}(t)$ are acceleration, velocity and displacement vectors, respectively, at time t, and $\mathbf{f}(t)$ is the excitation force vector. Assuming the mass matrix is known (Wang and Haldar, 1994) and the acceleration, velocity, and displacement response vectors are available, Eq. (1) can be rearranged as:

$$|\mathbf{u}(t)\ \dot{\mathbf{u}}(t)| \begin{Bmatrix} \mathbf{K} \\ \mathbf{C} \end{Bmatrix} = \{\mathbf{f}(t) - \mathbf{M}\ddot{\mathbf{u}}(t)\} \tag{2}$$

Equation (2) can be rewritten as:

$$\mathbf{A}(t) \cdot \mathbf{P} = \mathbf{F}(t); \quad \text{or,} \quad \mathbf{P} = \mathbf{A}(t)^* \cdot \mathbf{F}(t) \tag{3}$$

where **P** is a vector containing unknown stiffness and damping parameters to be identified and the matrix $\mathbf{A}(t)$ contains the measured dynamic responses of a structure. For measured $\mathbf{A}(t)$ and known $\mathbf{F}(t)$, the vector **P** can be estimated using an iterative procedure (Wang and Haldar, 1994).

In developing the ILS-UI procedure, the input excitation vector $\mathbf{f}(t)$ needs to be considered as unknown; thus the force vector $\mathbf{F}(t)$ defined in Eq. (3) is partially known and the system cannot be identified. Since the input excitation is not available, the iteration can be started by assuming it is zero at time t_i, $i = 1, 2, \ldots, p$, where $p \leq h$, and h is the total number of measured sample points. The algorithm is not sensitive to this initial assumption.

As the iteration process continues, the updated information on the system's unknown parameters, **P** emerges. Then, the information on the input excitation $\mathbf{f}(t)$ can be updated using Eq. (1) with the updated information on **P**. The iteration process continues until the convergence in the input excitation at all time points is obtained with a predetermined tolerance, ε. A tolerance of $\varepsilon = 10^{-4}$ is used in this study.

The procedure was extensively verified using noise-free and artificially noise-contaminated analytically generated response information.

3.2 *Modified iterative least squares with unknown input (MILS-UI) method*

The study leading to the development of the ILS-UI method established the validity of the concept, i.e., a system can be identified without excitation information. In developing ILS-UI, Wang and Haldar (1994) assumed viscous damping in Eq. (1). For a structure consisting of *ne* number of structural elements and using the ILS-UI method, the total number of parameters to be identified is 2 *ne* (*ne* numbers of stiffness and damping parameters). Since the changes in damping parameters of defective structure as compared to defect-free structure are not completely understood at present, the identified damping parameters cannot be utilized effectively for the defect identification

purpose. However, the presence of damping cannot be totally ignored in developing the dynamic governing equation. To increase the efficiency of the algorithm, particularly for large real structural systems, damping is considered to be Rayleigh-type; i.e., damping is proportional to the mass and stiffness (Ling and Haldar, 2004). For Rayleigh damping, matrix C in Eq. (1) becomes:

$$\mathrm{C} = \alpha\,\mathbf{M} + \beta\,\mathbf{K} \tag{4}$$

where α and β are the mass and stiffness proportional constants, respectively. These constants have close-form relationship with the first two natural frequencies (f_1 and f_2) of a structure (Clough and Penzien, 1993). Assuming the damping ratios in the first two modes are the same and for the known values of f_1 and f_2, α and β can be estimated. This is known as the modified ILS-UI or MILS-UI method (Ling and Haldar, 2004). Again, the procedure was extensively verified using analytically generated noise-free and noise-contaminated response information.

3.3 *Analytical verification of MILS-UI*

As will be discussed later, the MILS-UI method was experimentally verified by Vo and Haldar (2008a, b). Rectangular fixed ended and simply supported steel beams with uniform cross section: 3.81 cm wide, 0.64 cm thick, and 76.2 cm long, were tested in the laboratory. A typical test setup is shown in Figure 1. Before conducting the experiments, the beams were identified using the computer generated analytical response information. For the analytical study, the optimal number of finite elements of identical length required to represent the beams were found to be six (Vo and Haldar 2008a), as shown in Figure 2.

Since all the elements have the same length, the stiffness parameter for each element is expressed by the modulus of rigidity, EI, where E is the Young's modulus and I is the moment of inertia of the cross section of the beam. It is same for all the elements and is estimated to be 160.3 N-m^2.

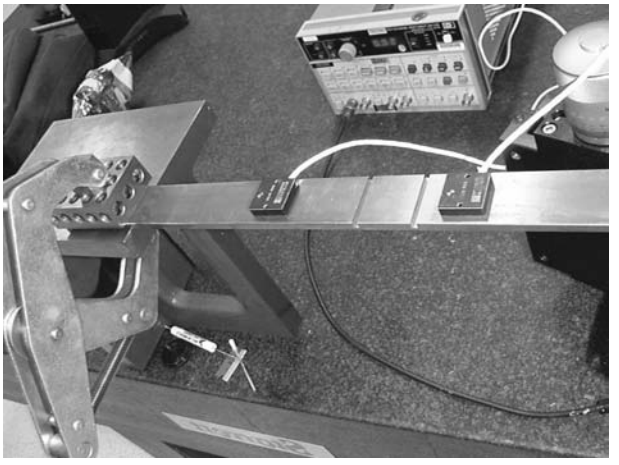


Figure 1 Typical experimental setup.

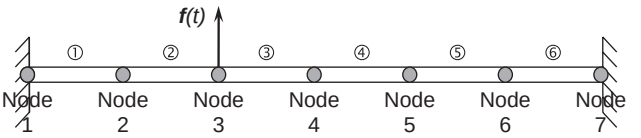


Figure 2 Finite element representation.

Table 1 Information on damping.

	ζ	f_1 (Hz)	f_2 (Hz)	α	β
Fixed	0.0160	54	145	7.84	2.47 E-5
Simply Supported	0.0175	25	99	3.73	3.69 E-5

Table 2 Identified *EI* values (N-m^2) for defect-free beams.

Element	*Fixed Ended Beam*	*Simply Supported Beam*
1	158.3	159.4
2	160.8	157.6
3	158.8	157.0
4	158.8	157.0
5	161.4	157.6
6	158.7	159.5

The amount of viscous damping ζ, expressed as the percentage of the critical for both the fixed and simply supported beams were obtained experimentally by the logarithmic decrement method (Clough and Penzien, 1993). Assuming the damping in the first two modes are the same and for the known values of the frequencies, the Rayleigh damping coefficients α and β can be estimated using standard procedures as discussed earlier. The information is summarized in Table 1.

3.3.1 *SHA of defect-free beams using analytical responses*

As mentioned earlier, required response information must be available to assess the health of a structure using the MILS-UI method. To obtain the analytical and experimental response information, the beam was excited by a sinusoidal force $f(t) = 0.15 \sin(100\pi t)$ applied at Node 3, located 25.4 cm from the left support, as shown in Figure 2. Analytical responses were calculated using a computer program for fixed ended and simply supported defect-free beams and the stiffness parameter (EI) for each of the six elements was identified using the MILS-UI method. The results are summarized in Table 2.

The expected EI value for all the elements is expected to be 160.3 N-m^2. The identified EI values for both beams are within about 2% of the expected value. Also, they are very similar indicating that the beam is defect-free. The results also imply that the MILS-UI method accurately identified the health of the defect-free beams.

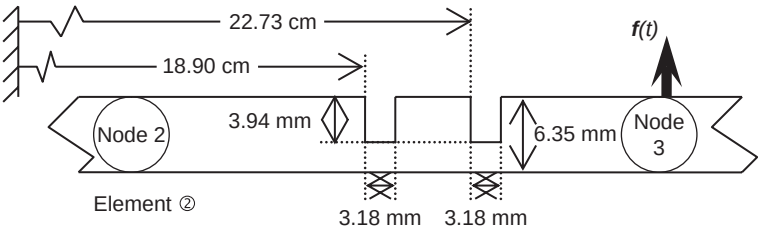


Figure 3 Defect sizes and locations in Element 2.

Table 3 Identified *EI* values (N-m^2) for defective beams.

Element	*Fixed Ended Beam*	*Simply Supported Beam*
1	99.5	24.9
2	74.1	20.2
3	150.6	49.9
4	150.6	50.4
5	148.6	51.1
6	153.3	52.4

3.3.2 *SHA of defective beams using analytical responses*

To study the capability of the method to detect defects, two cuts were introduced between Nodes 2 and 3 at 18.90 cm and 22.73 cm from the left support as shown in Figure 3. The notches are uniformly cut across the width of the beams' surface. Both notches have the same size, 3.2 mm wide and 3.94 mm deep.

For computer modeling, these notches are modeled by adding two nodes for each cut, one on each side of the defect. These four additional nodes for two notches create two defective elements of reduced thickness compared to the initial thickness of the beam. The responses at these additional nodes are not used to identify the defects. The responses of the defective fixed ended and simply supported beams are analytically obtained for the same sinusoidal load applied at Node 3.

For the defect identification purpose, the initial finite element arrangements were not changed since the presence of defects might not be known in advance. The stiffness parameter for each element was calculated using the MILS-UI method completely ignoring the excitation information. The results are summarized in Table 3.

Several important observations can be made from the results shown in Table 3. For the fixed ended beam, the stiffness parameter of Element 2 is reduced by about 54% from the expected value of 160.3 N-m^2. The reduction is the largest among all the elements indicating the defect may be in it. The stiffness parameter also is also reduced for Elements 1 and 3 to 6 but not as much. For the simply supported beam, the reductions of stiffness parameter are about 84% and 87% for Elements 1 and 2, respectively, and about 68% for the other 4 elements. For the simply supported case, the relative values of the modulus of rigidity and not the absolute values need to be considered for the defect detection. In any case, the stiffness parameters are

very different than the expected value indicating the presence of defect in the beam. The defect detection for the fixed ended beam is relatively simpler than the simply supported beam. The results establish the defect detection capability of the MILS-UI method.

3.4 *Experimental verification of MILS-UI*

After the successful verification using analytically obtained response information, the defect-free and defective beams were tested in the laboratory. A typical laboratory set up is shown in Figure 1 (Vo and Haldar, 2008b). Capacitive sensing element accelerometers made by Silicon Designs, Inc., Model 2210-005 and an autocollimator made by United Detector Technology, Model 431-XY were used to measure transverse and angular response time histories, respectively. A Hewlett-Packard function generator model 3314A was used to generate the voltage waveform. The waveform was fed through a midrange woofer speaker, Optimus model 40-1030, to convert electrical signal into displacements to excite the beams. The dynamic responses of the beams were recorded by a Tektronix model 2505 data logger with simultaneous sampling capability. Vo and Haldar (2008b) discussed in more detail the experiments.

After the placements of the accelerometers and the autocollimator, the beams were excited by the same sinusoidal excitation mentioned earlier and the responses were recorded simultaneously. A raw acceleration time history record may contain many sources of error including noise, high frequency content, slope and DC bias. These are conceptually shown in Figure 4. The responses are expected to be sinusoidal but plot shown in Figure 4 not even close to it. Since acceleration time histories are integrated to obtain velocity and displacement time histories required for the proposed method, it is important that all these errors are removed. Vo and Haldar (2003) discussed in great detail how to remove them efficiently and effectively. Post-processing of measured acceleration time history record is essential for the successful implementation of the proposed technique. This could be one of the major reasons for failure of other scholars to obtain satisfactory results using the SI-based approaches.

Error in numerical integration is another important element in post processing acceleration time histories. Several numerical integration methods such as the trapezoidal rule, Simpson's rule, and Boole's rule were considered by Vo and Haldar (2003).

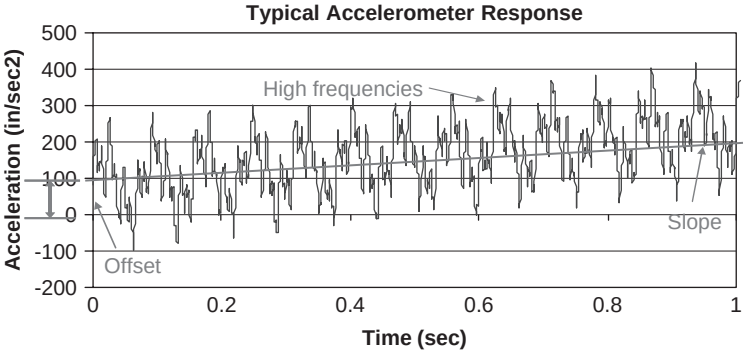


Figure 4 Typical recorded acceleration time history.

The mid-point rule and other open Newton-Cotes rules were not considered because they do not consider the end points. The consequence for not using the end points is the relative phase shift errors occurred as the time history being integrated. Vo and Haldar (2003) demonstrated that even though the trapezoidal rule produces the largest integration error compared to other methods, it is preferred over other rules for the problem under consideration because of its simplicity and efficiency in computing time.

3.4.1 *SHA of defect-free beams using measured responses*

Vo and Haldar (2008b) observed that simply post processing of the measured acceleration time histories may not identify the beams. Two additional sources of error need to be addressed. They are phase shift and amplitude errors. In short, the phase shift error can be mitigated by scaling all responses based on the measured response at a reference node. The amplitude error can be mitigated by using fewer nodal responses.

For the fixed ended and simply supported beams, since they were excited at Node 3, it is considered as the reference node as shown in Figure 5. Responses measured at Nodes 3 and 5 are used to identify the fixed ended beam. For the simply supported beam, the responses measured at Nodes 1, 3, 5, and 7 were used for the identification purpose. After measuring responses for the defect-free beams, the defect discussed in Section 3.2.2 were introduced and the responses were measured. The results are summarized in Table 4 for the fixed ended beam and in Table 5 for the simply supported beam.

The results in Table 4 indicate that for the subsection between Nodes 1 and 3, which includes Elements 1 and 2, the stiffness parameter (EI) is reduced by 36%, significantly more than the other two segments clearly indicating the location of the defects. To further pinpoint the location of the defects, the responses at Nodes 2 and 3 can be used. It was observed that the stiffness parameters for Element 1 and 2 were reduced by 19% and 39%, respectively, for this case. The results indicate that the defects are in Element 2. This exercise also indicates that the locations of the defects can be identified more accurately within a defective segment.

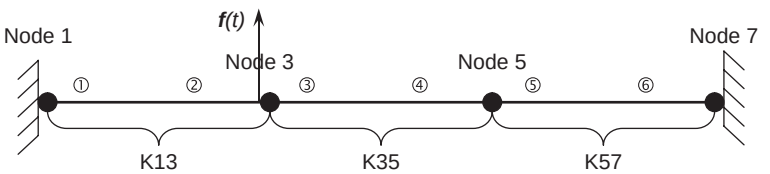


Figure 5 Identification of the fixed ended beam.

Table 4 Defect identification using measured responses for the fixed ended beam.

Beam Segment	*Undamaged EI (N-m²)*	*Undamaged EI (N-m²)*	*Percent Change*
K13	160.4	102.5	−36%
K35	186.7	185.3	−1%
K57	156.9	158.0	+1%

Table 5 Defect identification using measured responses for the simply supported beam.

Beam Segment	*Undamaged EI (N-m^2)*	*Undamaged EI (N-m^2)*	*Percent Change*
K13	143.6	22.7	−84%
K35	140.7	47.9	−66%
K57	143.9	48.4	−66%

Results shown in Table 5 for the simply supported beam indicate that the segment between Nodes 1 and 3 contains the defects. If the beam is identified using responses measured at Nodes 1, 2, 3, and 7, the rigidities for Elements 1 and 2 are reduced by 63% and 69%, respectively; indicating Element 2 contains the defects. As observed for the fixed ended beam, the results indicate that the location of defects can be identified more accurately if necessary.

4 The GILS-EKF-UI method

The discussions made in Section 3 clearly indicate that the health of structures can be assessed using only noise-contaminated response information, completely ignoring the information on excitation. However, the major weakness of the ILS-UI and MILS-UI methods is that the response information must be available at all dynamic degrees of freedom (DDOFs). This will not satisfy one of the major objectives of the new method discussed in Section 2. Kalman filter-based algorithm is generally used to identify a system if the response information is limited and noise-contaminated (Kalman, 1960). However, to implement the concept, the information on the input excitation and the initial state vector of the system to be identified must be available. Obviously, the information on them is not available. To assess the health of large structural systems, it is reasonable to assume that part of the structures will be instrumented. For a high rise building, the top two floors or the top two floors along with two floors at the middle of the building can be assumed to be instrumented. For the discussion purpose, the instrumented part(s) of the whole structure can be denoted as the substructure(s). The substructures are expected to satisfy all the requirements of the MILS-UI method and the stiffness and damping properties of all the elements in the substructures can be identified. Using the information judiciously, the initial estimate of the state vector for the whole structure can be obtained. It will be discussed in detail later. One of the byproducts of the MILS-UI method is that it will also identify the unknown excitation vector. This will satisfy the other major requirement of the Kalman filter-based algorithm. These observations prompted the research team to propose a two-stage approach by combining the MILS-UI and Kalman filter-based approaches producing the GILS-EKF-UI approach. The two stages of the approach are:

Stage 1 – Based on the available limited response information, select substructure(s) that will satisfy all the requirements of the MILS-UI method. Then identify the substructure(s) and generate information on the excitation.

Stage 2 – Since the information from Stage 1 will satisfy the Kalman filter-based algorithm, use it to identify the whole structure.

This way the whole structure can be identified with limited noise-contaminated response information producing the GILS-EKF-UI approach (Katkhuda and Haldar, 2007).

4.1 *Some of the important desirable features of the GILS-EKF-UI method*

The GILS-EKF-UI method will satisfy most of the desirable features discussed in Section 2. Some of its desirable features are discussed below.

1. Only parts of the structures need to be instrumented making it desirable from practical implementation point of view.
2. Since the excitation information need not be measured, instrumentation required to measure excitation information is expected to be relatively simple.
3. Since the procedure is finite element-based, the location(s) of the deteriorated elements can be identified with relative ease.
4. After repair and rehabilitation, the method can detect if the repair is done properly and if all the defects are repaired.
5. It will keep track of the rate of degradation if periodic assessments are conducted.
6. The procedure requires response information for a very short duration (about 0.2 sec). This will eliminate some of the sources of noise contamination in the measured responses.
7. Since the initial properties of the elements are selected from Stage 1 representing in-place instead of theoretical or assumed values, the mathematical model is expected to be more realistic avoiding the convergence issue, a major problem in the identification process.

4.2 *Salient features of GILS-EKF-UI*

The concepts behind the two stages of the GILS-EKF-UI method are briefly discussed below. They were discussed in more detail elsewhere (Katkhuda and Haldar, 2007).

Stage 1 – The mathematics of Stage 1 will be very similar to the discussions made in Section 3 for the MILS-UI method, except that Eq. (1) needs to be modified for the substructure. For Rayleigh-type damping, the governing equation for the substructure can be expressed as:

$$\mathbf{K}_{sub}\,\mathbf{x}_{sub}(t) + (\alpha\mathbf{M}_{sub} + \beta\mathbf{K}_{sub})\dot{\mathbf{x}}_{sub}(t) + \mathbf{M}_{sub}\,\ddot{\mathbf{x}}_{sub}(t) = \mathbf{f}_{sub}(t) \tag{5}$$

where $\mathbf{K}_{\text{sub}}$ and $\mathbf{M}_{\text{sub}}$ are the global stiffness and mass matrices, respectively, for the substructure, $\ddot{\mathbf{x}}_{sub}(t)$, $\dot{\mathbf{x}}_{sub}(t)$, and $\mathbf{x}_{sub}(t)$ are vectors containing acceleration, velocity and displacement at time t, respectively, α and β are Rayleigh damping coefficients discussed earlier, and $\mathbf{f}_{\text{sub}}(t)$ is unknown excitation force vector(s). $\mathbf{M}_{\text{sub}}$ and $\mathbf{K}_{\text{sub}}$ can be assembled from the mass and stiffness matrices of all the elements in the substructure by considering their connectivity and following the standard finite element formulation.

As mentioned earlier, the mass matrix is generally assumed to be known. Suppose, $\mathbf{K}^i$ denotes the stiffness matrix of the two-dimensional ith beam of uniform cross section. It can be represented as:

$$\mathbf{K}^i = (E_i I_i / L_i)[f(A_i,\, I_i,\, L_i]_{6\times 6} = k_i \mathbf{S}^i \tag{6}$$

where E_i, I_i, and A_i are Young's modulus, moment of inertia, and area of the cross section of the ith beam element, respectively, $k_i = E_i I_i / L_i$, and $\mathbf{S}^{\mathrm{i}}$ is a matrix of size 6×6. $\mathbf{K}_{\mathrm{sub}}$ can be expressed as:

$$\mathbf{K}_{sub} = \sum_{i=1}^{nesub} k_i \mathbf{S}^{\mathrm{i}} = k_1 \mathbf{S}^1 + k_2 \mathbf{S}^2 + \cdots + k_{nesub}\, \mathbf{S}^{nesub} \tag{7}$$

where, *nesub* is the total number of elements in the substructure. In Stage 1, the **P** vector in Eq. (3) contains all the unknown parameters in the substructure to be identified and can be shown to be:

$$\mathbf{P} = [k_1,\, k_2, \ldots,\, k_{nesub}, \beta k_1,\, \beta k_2, \ldots, \beta k_{nesub},\, \alpha]^T \tag{8}$$

All the parameters were defined earlier. All the necessary information to implement the MILS-UI method for the substructure is now available. At the completion of Stage 1, the time history of the unknown excitation force, the Rayleigh-damping coefficients and the stiffness parameters of all the elements in the substructure will be available. The information on damping will be applicable to the whole structure.

Stage 2 – In Stage 1, the mathematical model used is considered to be perfect and the measured response information is considered to be noise-free. The imperfect nature of mathematical modeling and noise in the measured responses need to be incorporated in Stage 2, and the Extended Kalman Filter (EKF) method will be ideal to address them. In order to apply the EKF method, the state vector can be defined as:

$$\mathbf{Z}(t) = \begin{bmatrix} \mathbf{Z}_1(t) \\ \mathbf{Z}_2(t) \\ \mathbf{Z}_3(t) \end{bmatrix} = \begin{bmatrix} \mathbf{X}(t) \\ \dot{\mathbf{X}}(t) \\ \tilde{\mathbf{K}} \end{bmatrix} \tag{9}$$

where $\mathbf{Z}(t)$ is the state vector at time t, $\mathbf{X}(t)$ and $\dot{\mathbf{X}}(t)$ are the displacement and velocity vectors, respectively, at time t for the whole structure, and $\tilde{\mathbf{K}}$ is a vector containing the element stiffness parameters of the whole structure that need to be identified. $\tilde{\mathbf{K}}$ can be shown to be:

$$\tilde{\mathbf{K}} = [k_1 \;\; k_2 \;\; \cdots \;\; k_{ne}]^{\mathrm{T}} \tag{10}$$

where *ne* is the total number of elements in the whole structure. It is assumed that the stiffness will not change with time during the identification process.

The equation of motion can be expressed in a state equation as:

$$\dot{\mathbf{Z}}(t) = \begin{bmatrix} \dot{\mathbf{Z}}_1(t) \\ \dot{\mathbf{Z}}_2(t) \\ \dot{\mathbf{Z}}_3(t) \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{X}}(t) \\ \ddot{\mathbf{X}}(t) \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{X}}(t) \\ -\mathbf{M}^{-1}(\mathbf{K}\mathbf{X}(t) + (\alpha\mathbf{M} + \beta\mathbf{K})\dot{\mathbf{X}}(t) - \mathbf{f}(t)) \\ 0 \end{bmatrix} \tag{11}$$

where **K** and **M** are the global stiffness and mass matrixes of the whole structure, respectively, and f(t) is the input excitation force vector identified in Stage 1.

Equation (11) can be mathematically expressed as:

$$\frac{d\,\mathbf{Z}}{dt} = \dot{Z}(t) = f[\mathbf{Z}(t),\, t] \tag{12}$$

The initial state vector $\mathbf{Z}_0$ is assumed to be Gaussian with a mean vector of $\hat{\mathbf{Z}}_0$ and a constant error covariance matrix of **D**. It is generally denoted as $\mathbf{Z}_0 \sim N(\hat{\mathbf{Z}}_0, \mathbf{D})$.

Suppose the response of the structure is measured at time t_k. The observational vector $\mathbf{Y}_{t_k}$ of size $(B \times 1)$, where B is the total number of displacement and velocity observations (the information of acceleration is not required for the second stage), can be expressed as:

$$\mathbf{Y}_{t_k} = \mathbf{H}\mathbf{Z}(t_k) + \mathbf{V}_{t_k} \tag{13}$$

where $\mathbf{Z}(t_k)$ is the state vector of size $(2\,N + L) \times 1$ at time t_k, N is the total number of DDOFs, L is the total number of unknown stiffness parameters, **H** is a matrix of size $[B \times (2N + L)]$ containing information of measured responses, and $\mathbf{V}_{t_k}$ is the observational noise vector of size $(B \times 1)$, assumed to be Gaussian white noise with zero mean and a covariance of $\mathbf{R}_{t_k}$. It is generally denoted as $\mathbf{V}_{t_k} \sim N(0, \mathbf{R}_{t_k})$.

Stage 2 can be carried out in the following steps:

Step 1. Define the initial state vector $\hat{\mathbf{Z}}_0(t_0/t_0)$ and its error covariance $\mathbf{D}(t_0/t_0)$. The initial state vector can be expressed as:

$$\hat{\mathbf{Z}}_0(t) = \begin{bmatrix} \mathbf{Z}_1(t_0/t_0) \\ \mathbf{Z}_2(t_0/t_0) \\ \mathbf{Z}_3(t_0/t_0) \end{bmatrix} = \begin{bmatrix} \mathbf{X}(t_0/t_0) \\ \dot{\mathbf{X}}(t_0/t_0) \\ \tilde{\mathbf{K}}(t_0/t_0) \end{bmatrix} \tag{14}$$

where $\mathbf{X}(t_0/t_0)$ and $\dot{\mathbf{X}}(t_0/t_0)$ are the displacement and velocity vectors, respectively, and $\mathbf{X}(t_0/t_0) = \mathbf{X}(t)$ and $\dot{\mathbf{X}}(t_0/t_0) = \dot{\mathbf{X}}(t)$. As discussed earlier, only acceleration responses will be measured at a few DDOFs to implement the procedure. The acceleration time histories will be successively integrated to obtain the required velocity and displacement time histories. The initial parts of acceleration, velocity, and displacement time histories, say for the duration between 0.0 and 0.05 sec were not used for the identification purpose to avoid the effect of initial boundary conditions and the integration error.

The initial value of the stiffness parameters for all the elements in the structure $\tilde{\mathbf{K}}(t_0/t_0)$ can be assumed to be:

$$\tilde{\mathbf{K}}(t_0/t_0) = \begin{bmatrix} k_1(t_0/t_0) \\ k_2(t_0/t_0) \\ k_3(t_0/t_0) \\ \vdots \\ k_{ne}(t_0/t_0) \end{bmatrix} = \begin{bmatrix} k_i \\ k_i \\ k_i \\ \vdots \\ k_i \end{bmatrix} \tag{15}$$

Assuming the substructure in Stage 1 has only two elements, one beam and another column, and their identified stiffness parameters are k_1 and k_2, then $k_i(t_0/t_0)$ in Eq. (15) can be assumed as k_1 or k_2 depending upon whether the element under consideration is a beam or a column. Conceptually, any initial values of k_i's can be assumed. However, if the initial assumed values are far from the actual values, it may create a convergence problem or reduce the efficiency of the algorithm. Stiffness parameters of all the elements in the similar category in a real structure are expected to be similar, thus this assumption is quite reasonable.

The initial error covariance matrix $\mathbf{D}(t_0/t_0)$ contains information on the errors in the velocity and displacement responses and in the initial estimate of the stiffness parameters of the elements. It is generally assumed to be a diagonal matrix and can be expressed as (Hoshiya and Saito, 1984; Wang and Haldar, 1997):

$$\mathbf{D}(t_0/t_0) = \begin{bmatrix} \mathbf{D}_x(t_0/t_0) & 0 \\ 0 & \mathbf{D}_k(t_0/t_0) \end{bmatrix} \tag{16}$$

where $\mathbf{D}_x(t_0/t_0)$ is a matrix of size $(2\,N \times 2\,N)$ and N is the total number of DDOFs in the whole structure. This represents the initial errors in the velocity and displacement responses, and is assumed to have a value of 1.0 in the diagonals. $\mathbf{D}_k(t_0/t_0)$ is a diagonal matrix of size $(L \times L)$ and L is the total number of the unknown stiffness parameters to be identified, as mentioned earlier. It contains the initial covariance matrix of $\tilde{\mathbf{K}}$. Hoshiya and Saito (1984) and Jazwinski (1970) pointed out that the diagonals should be large positive numbers to accelerate the convergence of the local iteration process, and 1000 can be used for this purpose. The same value is used in this study.

Step 2. *Prediction phase*: In the context of EKF, the predicted state $\hat{\mathbf{Z}}(t_{k+1}/t_k)$ and its error covariance $\mathbf{D}(t_{k+1}/t_k)$ are evaluated by linearizing the nonlinear dynamic equation as:

$$\hat{\mathbf{Z}}(t_{k+1}/t_k) = \hat{\mathbf{Z}}\,(t_k/t_k) + \int_{t_k}^{t_{k+1}} f[\hat{\mathbf{Z}}(t/t_k), t]dt \tag{17}$$

and

$$\mathbf{D}(t_{k+1}/t_k) = \mathbf{\Phi}[t_{k+1}, t_k; \hat{\mathbf{Z}}(t_k/t_k)] \bullet \mathbf{D}(t_k/t_k) \bullet \mathbf{\Phi}^{\mathrm{T}}[t_{k+1}, t_k; \hat{\mathbf{Z}}(t_k/t_k)] \tag{18}$$

where $\mathbf{\Phi}[t_{k+1}, t_k; \hat{\mathbf{X}}(t_k/t_k)]$ is the state transfer matrix from time t_k to t_{k+1} and can be written in an approximate form as:

$$\mathbf{\Phi}[t_{k+1}, t_k; \hat{\mathbf{Z}}(t_k/t_k)] = \mathbf{I} + \Delta t \bullet \mathbf{F}[t_k; \hat{\mathbf{Z}}(t_k/t_k)] \tag{19}$$

In which, $\mathbf{I}$ is a unit matrix and

$$\mathbf{F}[t_k; \hat{\mathbf{Z}}(t_k/t_k)] = \left[\frac{\partial f\,[\mathbf{Z}(t_k), t_k]}{\partial Z_j}\right]_{\mathbf{Z}(t_k)=\hat{\mathbf{Z}}(t_k/t_k)} \tag{20}$$

where Z_j is the jth component of the vector $\mathbf{Z}(t_k)$.

Step 3. *Updating phase*: Since observations are available at $k+1$, the state vector and the error covariance matrix can be updated as:

$$\begin{aligned}\hat{\mathbf{Z}}(t_{k+1}/t_{k+1}) = {} & \hat{\mathbf{Z}}(t_{k+1}/t_k) + \mathbf{K}[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)] \bullet \{\mathbf{Y}(t_{k+1}) \\ & - \mathbf{H} \bullet [\hat{\mathbf{Z}}(t_{k+1}/t_k), t_{k+1}]\}\end{aligned} \tag{21}$$

$$\begin{aligned}\mathbf{D}(t_{k+1}/t_{k+1}) = {} & \left\{\mathbf{I} - \mathbf{K}[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)] \bullet \mathbf{M}[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)\right\} \\ & \bullet \mathbf{D}(t_{k+1}/t_k) \bullet \{\mathbf{I} - \mathbf{K}[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)] \bullet \mathbf{M}[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)]\}^T \\ & + \mathbf{K}[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)] \bullet \mathbf{R}(t_{k+1}) \bullet \mathbf{K}^T[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)]\end{aligned} \tag{22}$$

where $\mathbf{K}\,[t_{k+1}; \hat{\mathbf{Z}}(t_{k+1}/t_k)]$ is the Kalman gain matrix.

Step 4. By taking the next time increment, i.e., $k = k+1$ and using Eqs. (17), (18), (21), and (22), the system parameters are updated. This procedure will continue until all the time points are used, i.e., $k = h$, where h represents the total number of discrete time points of the measurements.

The iteration process covering all the time points is generally defined as the local iteration. When the local iteration procedure is completed, Hoshiya and Saito (1984) suggested incorporating a weighted global iterative procedure with an objective function into the local iteration to obtain the stable and convergent estimation of the parameters to be identified.

To start the first global iteration, the initial values of $\hat{\mathbf{Z}}^{(1)}(t_h/t_h)$ and $\mathbf{D}^{(1)}(t_h/t_h)$ need to be assumed, or the information from Step 4 stated above can be used to increase the efficiency, where superscript (1) represents the first global iteration, and they can be expressed as:

$$\hat{\mathbf{Z}}^{(1)}(t_h/t_h) = \begin{bmatrix}\mathbf{Z}_1^{(1)}(t_0/t_0) \\ \mathbf{Z}_2^{(1)}(t_0/t_0) \\ \mathbf{Z}_3^{(1)}(t_h/t_h)\end{bmatrix} = \begin{bmatrix}X^{(1)}(t_0/t_0) \\ \dot{X}^{(1)}(t_0/t_0) \\ \tilde{\mathbf{K}}^{(1)}(t_h/t_h)\end{bmatrix} \tag{23}$$

and

$$\mathbf{D}^{(1)}(t_h/t_h) = \begin{bmatrix}\mathbf{D}_x(t_0/t_0) & 0 \\ 0 & \mathbf{D}_k(t_h/t_h)\end{bmatrix} \tag{24}$$

In the second global iteration, a weight factor w is introduced to the error covariance matrix to accelerate the local EKF iteration. Hoshiya and Saito (1984) observed that to obtain better and stable convergence, the value of w should be a large positive number. Hoshiya and Saito (1984), Koh, See, and Balendra (1995), and Oreta and Tanabe (1994) assumed w to be 100 and Hoshiya and Sutoh (1993) assumed it to be in the range of 1000 to 10000 in some applications. In this study, w is considered to be 100.

To start the second global iteration, the initial values of the state vector $\hat{\mathbf{Z}}^{(2)}(t_0/t_0)$ and the error covariance matrix $\mathbf{D}^{(2)}(t_0/t_0)$ can be shown to be:

$$\hat{\mathbf{Z}}^{(2)}(t_0/t_0) = \hat{\mathbf{Z}}^{(1)}(t_b/t_b) = \begin{bmatrix} \hat{\mathbf{Z}}_1^{(2)}(t_0/t_0) \\ \hat{\mathbf{Z}}_2^{(2)}(t_0/t_0) \\ \hat{\mathbf{Z}}_3^{(2)}(t_0/t_0) \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(2)}(t_0/t_0) \\ \dot{\mathbf{X}}^{(2)}(t_0/t_0) \\ \tilde{\mathbf{K}}^{(2)}(t_0/t_0) \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)}(t_0/t_0) \\ \dot{\mathbf{X}}^{(1)}(t_0/t_0) \\ \tilde{\mathbf{K}}^{(1)}(t_b/t_b) \end{bmatrix} \tag{25}$$

and

$$\mathbf{D}^{(2)}(t_0/t_0) = \begin{bmatrix} \mathbf{D}_x(t_0/t_0) & 0 \\ 0 & w\,\mathbf{D}_k^{(1)}(t_b/t_b) \end{bmatrix} \tag{26}$$

where w is the weight factor used to accelerate the local iteration, as discussed earlier and $\mathbf{D}_k^{(1)}(t_b/t_b)$ is the error covariance matrix corresponding to the parameter $\tilde{\mathbf{K}}$ at time t_b in the last global iteration. With this information, the prediction and updating phases of the local iteration are carried out for all the time points, producing the state vector and the error covariance matrix for the second global iteration. The information can then be used to initiate the third global iteration. The global iterations are repeated until a predetermined convergence criterion is satisfied, i.e., $|\hat{\mathbf{Z}}^i(t_b/t_b) - \hat{\mathbf{Z}}^{i-1}(t_b/t_b)| \leq \varepsilon$, where i represents the iteration number and ε is the tolerance level to be used in the numerical evaluation of the stiffness parameters. Since the stiffness parameters are of the order of 100 in this study, ε is considered to be 0.1.

5 Experimental verification of GILS-EKF-UI

After the successful completion of the beam experiments, comprehensive analytical and experimental verification programs were undertaken to verify ILS-UI, MILS-UI, and GILS-EKF-UI methods for frame structures. Before conducting any experiment, the particular situation was thoroughly investigated analytically (Katkhuda and Haldar, 2006; Katkhuda, Martinez-Flores, and Haldar, 2005). Only, experimental investigation involving a steel frame is emphasized in the following discussions. It is to be noted that the measured responses are expected to contain several sources of noise.

5.1 *Test specimen*

Details of the experiment can be found elsewhere (Martinez-Flores and Haldar 2007; Martinez-Flores, Katkhuda, and Haldar, 2008). Only essential features are discussed briefly below. A scaled two-dimensional one-bay three-story steel frame, shown pictorially in Figure 6, was built in the laboratory. The width and story height of the frame were 3.05 m and 1.22 m, respectively. An I-section with nominal cross-sectional area

Figure 6 Test specimen.

Figure 7 Reconfigurable connections.

and moment of inertia of 14.58 cm^2 and 253.9 cm^4, respectively, produced from the same batch, was used for all beams and columns. The two supports were considered to be fixed.

Bolted connections were used so that the frame can be reconfigured as shown in Figure 7. Before conducting any experiment, the amount of damping present is the system was estimated experimentally. The log decrement method was used for this

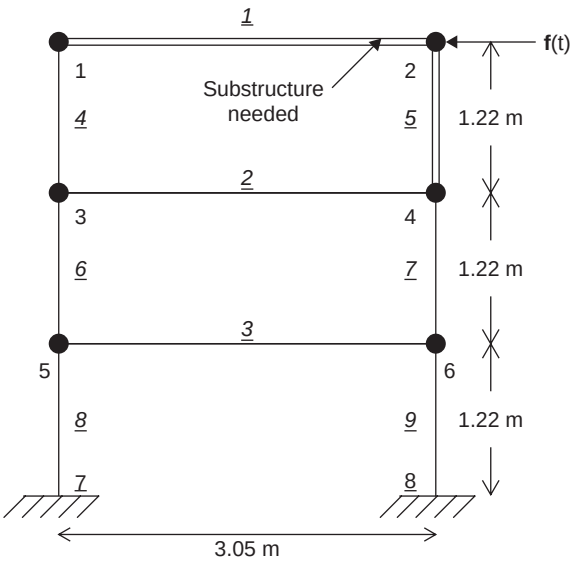


Figure 8 Finite element representation of the frame.

purpose (Clough and Penzien, 1993). Then, assuming the same damping for the first two significant frequencies, a procedure suggested by Clough and Penzien (1993) was used to estimate the Rayleigh damping coefficients α and β. For the defect-free frame, the equivalent modal damping is found to be 0.01201 and the corresponding α and β are estimated to be 0.96337 and 7.3261E-05, respectively.

The frame is represented by three beams and six columns as shown in Figure 8. The stiffness of each element is expressed in terms of a stiffness parameter, (EI/L). The nominal stiffness parameter for all the beams and columns are 96.5 kN-m and 242.2 kN-m, respectively. The nominal mass for all the elements was estimated to be 11.2 kg/m.

Assuming the bases are fixed, the frame can be represented by 18 DDOFs. The first two natural frequencies of the frame were estimated to be 9.76 Hz and 34.12 Hz, respectively. To obtain response information, the frame was excited by a harmonic force of $\mathbf{f}(t) = 0.0014 \sin(58.23t)$ kN, applied horizontally at the top floor at Node 2, as shown in Figure 8.

5.2 *Identification of defect-free frame with GILS-EKF-UI*

As mentioned earlier, to apply the GILS-EKF-UI method to identify the frame, a substructure is necessary. The substructure shown in the study is shown in Figure 8. Using responses measured at Nodes 1, 2, and 4, with 9 DDOFs, the stiffness parameter of Elements 1 and 5 are identified to be 93.5 kN-m and 235.5 kN-m with an error in identification of 2.98% and 2.63%, respectively. With the information on the excitation force and the initial state vector, the whole frame was identified. The results are summarized in Table 6.

Table 6 Stiffness parameter identification for the defect-free frame.

Element	*Nominal Value (kN-m)*	*Identified Stiffness Parameter (kN-m)*		
		MILS-UI		*GILS-EKF-UI*
(1)	*(2)*	*(3)* Analy.	*(4)* Exp.	*(4)* Exp.
1	96.5	96.5	98.9	93.2
2	96.5	97.3	99.6	93.5
3	96.5	97.2	99.8	93.5
4	242.2	243.5	250.1	236.1
5	242.2	243.4	250.1	234.7
6	242.2	243.5	250.3	234.6
7	242.2	243.5	250.3	234.2
8	242.2	243.3	250.3	235.8
9	242.2	243.4	250.3	235.1

For comparison, the identified stiffness parameter using the MILS-UI method using the analytical and experimental responses with 18 DDOFs are also given in the table. The corresponding maximum errors in the identifications are about 0.8% and 3.3%, respectively. The results are expected. The analytical responses, in the absence of any noise, are expected to identify the structure more accurately. When the GILS-EKF-UI was used, the maximum error was observed to be about −2.7%. In this case, the stiffness parameter values for all the elements went down. However, the reductions are very similar indicating the frame is defect-free.

5.3 *Defect identification with GILS-EKF-UI*

To evaluate the capability of the GILS-EKF-UI method to detect defects, several defects were introduced in the frame (Haldar and Martinez-Flores, 2008). Responses were measured in the presence of the defects. Presences of defects were assumed to be unknown at the time of investigation. Thus, without changing the initial defect-free finite element representation, the stiffness parameters in the presence of defects were identified. Some of the defects considered in the study are:

Case 1 – Removal of beam 3 in Figure 8 assuming that it is completely broken
Case 2 – Reduction of cross sectional area of a beam over a finite length
Case 3 – Presence of multiple cracks in a beam
Case 4 – Presence of single crack in a beam

5.3.1 *Removal of Beam 3*

A beam represented by Element 3 in Figure 8 is removed completely from the frame, representing the element has suffered a severe amount of damage. The stiffness for the beam is supposed to be zero. However, since it was not known in advance for the defect detection purpose, the initial finite element representation was not changed.

Table 7 Stiffness parameter identification with one beam removed.

Element	*Nominal EI/L (kN-m)*	*Identified (EI/L) values in (kN-m)*	
		9 DDOFs	*10 DDOFs*
1	96.5	84.4	89.2
2	96.5	83.8	89.4
3	**96.5**	**−2.1**	**−1.3**
4	242.2	220.5	225.6
5	242.2	221.3	226.8
6	242.2	219.8	224.2
7	242.2	219.7	222.9
8	242.2	220.1	226.4
9	242.2	219.8	226.3

For this particular defect, the Rayleigh damping coefficients α and β are estimated to be 1.5502 and 1.83E-04, respectively. The responses recorded at 0.00025 sec time intervals from 0.05 to 0.9 sec providing 3401 time points and considering the same substructure shown in Figure 8, the stiffness parameters of Elements 1 and 5 identified to be 90.3 kN-m and 224.8 kN-m, respectively. Then using responses at Nodes 1, 2, and 4, the whole frame was identified and the results are shown in Table 7.

The results show that the stiffness parameters for all the elements are similar to what were expected. However, the stiffness parameter of Element 3 is found to be very small, close to zero and negative, indicating the location of the defect.

To study the effect of additional response information in identifying the frame, an additional horizontal time history recorded at node 3 was added to the 9 response time histories just discussed. Thus, using 10 responses, the frame was identified. As expected, the identified stiffness parameter for Element 3 approaches zero more closely when more response information is used, indicating the beneficial effect of additional information.

5.3.2 *Reduction of cross sectional area over a finite length*

In this defective scenario, loss of area is simulated by reducing the cross sectional area over a limited length of a beam (Element 3) in Figure 8. The area is removed using a vertical milling machine. This defect is pictorially shown in Figure 9. Its analytical representation is shown in Figure 10.

For theoretical verification, one additional element of length equal to 5.08 cm located 1.2 m from the left hand column is introduced. The defect represents a 45.6% reduction in the cross sectional area and 46.4% reduction in the moment of inertia. The equivalent modal damping ζ is experimentally found to be 0.01402 for this case. The corresponding Rayleigh damping coefficients α and β are found to be 0.93732 and 7.913E-05, respectively.

As before, using the substructure shown in Figure 8, the stiffness parameters for all the elements were identified assuming the responses are available at 9 DDOFs and 10

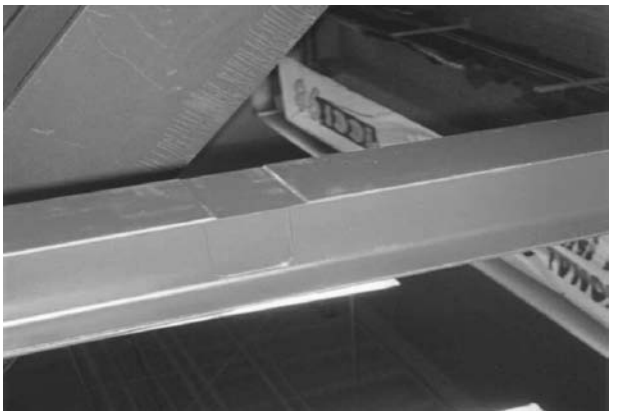


Figure 9 Reduction in area over a finite length.

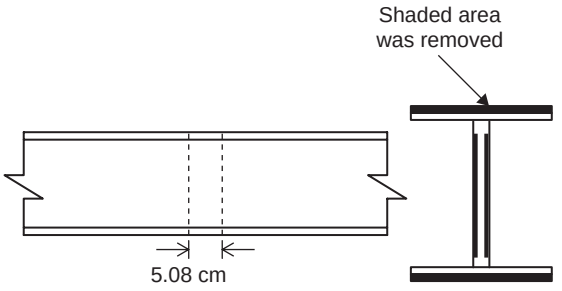


Figure 10 Analytical representation of reduction in area.

Table 8 Stiffness parameter identification with reduced area.

Elem	*Nominal EI/L (kN-m)*	*Identified (EI/L) values in (kN-m)*			
		9 DDOFs	*% change*	*10 DDOFs*	*% change*
1	96.5	90.4	−6.3	92.6	−4.0
2	96.5	90.0	−6.7	92.5	−4.1
3	**96.5**	**87.6**	**−9.2**	**86.9**	**−9.9**
4	242.2	229.3	−5.3	234.6	−3.1
5	242.2	228.9	−5.4	235.1	−2.9
6	242.2	230.9	−4.6	231.8	−4.3
7	242.2	231.9	−4.0	233.1	−3.8
8	242.2	229.3	−5.3	234.6	−3.1
9	242.2	229.6	−5.2	234.9	−3.0

DDOFs. The results are summarized in Table 8. In this case also, the stiffness parameters for all the elements except for Element 3 are similar to what are expected. However, the stiffness parameter for Element 3 is reduced by the largest amount indicating that it contains the defect.

Figure 11 Four cracks in a beam.

Table 9 Stiffness parameter identification with multiple cracks.

Element	*Nominal EI/L (kN-m)*	*Identified (EI/L) (kN-m)*	*% change*
1	96.5	89.4	−7.4
2	96.5	89.6	−7.2
3	**96.5**	**71.5**	**−25.9**
4	242.2	223.4	−7.8
5	242.2	221.9	−8.4
6	242.2	222.5	−8.1
7	242.2	223.1	−7.9
8	242.2	224.7	−7.2
9	242.2	224.8	−7.2

5.3.3 *Presence of multiple cracks in a beam*

In this defect scenario, four saw cut notches of width of 1.5 mm and depth of 9.1 cm each were introduced in beam 3 as pictorially shown in Figure 11. They are spaced 30 cm center to center from node 5.

As before, using the same substructure and without changing the initial finite element representation, the whole frame was identified. The identified stiffness parameters for all the elements are summarized in Table 9 using responses available at 10 DDOFS.

In this case, the stiffness parameter for Element 3 is reduced by the largest amount. The stiffness parameters for all other elements are similar to what are expected. The results indicate that Element 3 is defective.

5.3.4 *Presence of a crack in a beam*

In this defect scenario, a single saw cut notch is introduced as shown in Figure 12. For theoretical verification, one additional element of length equal to 1.5 cm located 1.2 m from the left hand column is introduced, as shown in Figure 13. The defect represents

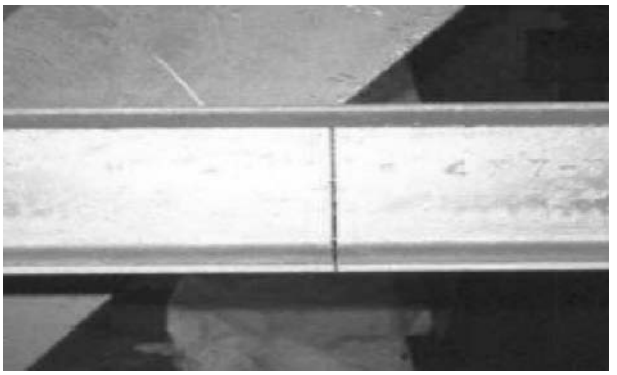


Figure 12 Presence of a crack.

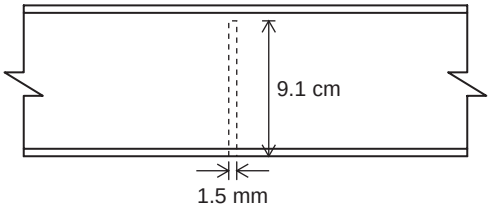


Figure 13 Analytical representation of the crack.

Table 10 Stiffness parameter identification with one crack.

Element	*Nominal EI/L (kN-m)*	*Identified (EI/L) (kN-m)*	*% change*
1	96.5	92.3	−4.4
2	96.5	92.7	−3.9
3	**96.5**	**88.7**	**−8.1**
4	242.2	233.6	−3.6
5	242.2	232.3	−4.1
6	242.2	233.2	−3.7
7	242.2	233.8	−3.5
8	242.2	233.3	−3.7
9	242.2	233.9	−3.4

a 63.9% reduction in the cross sectional area and 78.0% reduction in the moment of inertia. In this case, the equivalent modal damping ζ is experimentally found to be 0.01282. The corresponding Rayleigh damping coefficients α and β are found to be 0.94230 and 7.79E-05, respectively.

As before, using the same substructure and without changing the initial finite element representation, the whole frame was identified. The identified stiffness parameters for all the elements are summarized in Table 10 using responses available at 10 DDOFS.

In this case also, the reduction is the stiffness parameter for Element 3 is the largest, indicating it contains the defect.

6 Conclusions

A novel structural health assessment technique, known as the GILS-EKF-UI method, is presented. The unique feature of the method is that it can identify defects at the element level using only minimum noise-contaminated dynamic response information. The method is verified using both analytically and experimentally obtained dynamic response information. For laboratory experiments, a two-dimensional three story one-bay steel frame built to 1/3 scale was tested. After successfully identifying the defect-free frame using the GILS-EKF-UI method, several defects were introduced in the frame. Four of them, i.e., removing a member, reducing the cross sectional area over a finite length, multiple cracks in a member, and a single crack in a member are discussed in the paper. In all cases, the method correctly identified the location of defects. As expected, the defect detection potential increases if the defect is relatively large. If the measured responses are post-processed properly, the presence of noise in the data is not a factor in identifying defects. It is a significant development in the health assessment of real structures since the concept was found to be unsuitable for the defect identification purpose in the late seventies. It can be considered as a nondestructive defect evaluation technique for existing structures on a continuing basis or just after major natural or man-made events.

References

Clough, R.W. & Penzien, J., *Dynamics of Structures*. 2nd Ed., McGraw-Hill, Inc., 1993.

Haldar, A. & Martinez-Flores, R., A Novel Methodology for Assessment of Safety of Existing Structures. *Journal of Structural Engineering*, Vol. 35, No. 1, pp. 82–91, 2008.

Haldar, A., Martinez-Flores, R. & Katkhuda, H., Crack Detection in Existing Structures Using Noise-Contaminated Dynamic Responses. *Theoretical Applied Fracture Mechchanics*, doi:10.1016/j.tafmec.2008.04.007, 2008 (in press).

Hoshiya, M. & Saito, E., Structural Identification by Extended Kalman Filter. *Journal of Engineering Mechanics, ASCE*, Vol. 110, No. 12, pp. 1757–1770, 1984.

Hoshiya, M. & Sutoh, A., Kalman Filter- Finite Element Method in Identification. *Journal of Engineering Mechanics, ASCE*, V. 119, No. 2, pp. 197–210, 1993.

Jazwinski, A.H., *Stochastic Process and Filtering Theory.* Academic Press, 1970.

Kalman, R.E., A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Transaction of the ASME*, pp. 35–45, 1960.

Katkhuda, H. & Haldar, A., A Novel Health Assessment Technique with Minimum Information. *Structural Control and Health Monitoring*, published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/stc.221, 2007.

Katkhuda, H. & Haldar, A., Defect Identification under Uncertain Blast Loading. *Optimization and Engineering Journal,* Vol. 7, No. 3, pp. 277–296, 2006.

Katkhuda, H., Martinez-Flores, R. & Haldar, A., Health Assessment at Local Level With Unknown Input Excitation. *Journal of the Structural Engineering, ASCE,* Vol. 131, No. 6, pp. 956–965, June, 2005.

Koh, C.G., See, L.M. & Balendra, T., Damage Detection of Buildings: Numerical and Experimental Studies. *Journal of Structural Engineering, ASCE*, Vol. 121, No. 8, pp. 1155–60, 1995.

Ling, X. & Haldar, A., Element Level System Identification with Unknown Input with Rayleigh Damping. *Journal of Engineering Mechanics, ASCE*, Vol. 130, No. 8, pp. 877–885, 2004.

Martinez-Flores, R., Katkhuda, H. & Haldar, A., A Novel Health Assessment Technique with Minimum Information: Verification. *International Journal of Performability Engineering,* Vol. 4, No. 2, pp. 121–140, 2008.

Martinez-Flores, R. & Haldar, A., Experimental Verification of a Structural Health Assessment Method without Excitation Information. *Journal Of Structural Engineering*, Vol. 34, No. 1, pp. 33–39, April-May, 2007.

Martinez-Flores, R., Haldar, A. & Katkhuda, H., Structural Health Assessment After an Impact. *Paper No. IMECE 2006-13718, ASME,* 2006.

Maybeck, P., *Stochastic Models, Estimation, and Control*, Academic Press. U.K., 1979.

Mehrabian, A. & Haldar, A., Seismic Performance of Steel Frames with a Post-Northridge Connection, *Report No. CEEM-02-001*, Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, Arizona, 2002.

Minnesota Department of Transportation Metro District, Fracture Critical Bridge Inspection, Bridge #9340 (Squirt Bridge), June 2006.

Oreta, A. & Tanabe, T., Element Identification of Member Properties of Framed Structures. *Journal of Structural Engineering, ASCE*, V. 20, No. 7, pp. 1961–1976, 1994.

Vo, P.H. & Haldar, A., Health Assessment of Beams – Theoretical Formulation and Analytical Verification. *Structure and Infrastructure Engineering*, Vol. 4, No. 1, pp. 33–44, 2008a.

Vo, P.H. & Haldar, A., Health Assessment of Beams – Experimental Verification. *Structure and Infrastructure Engineering*, Vol. 4, No. 1, pp. 45–56, 2008b.

Vo, P.H. & Haldar, A., Post Processing of Linear Accelerometer Data in System Identification. *Journal of Structural Engineering*, Vol. 30, No. 2, pp. 123–130, 2003.

Wang, & Haldar, A., System Identification with Limited Observations and Without Input. *J. of Engineering Mechanics, ASCE*, 123(5) (1997) 504–511.

Wang, D. & Haldar, A., An Element Level SI with Unknown Input Information. *Journal of Engineering Mechanics Division, ASCE*, Vol. 120, No. 1, pp. 159–176, 1994.

Zhao, Z. & Haldar, A., Fatigue Damage Evaluation and Updating Using Nondestructive Inspections. *Journal of Engineering Fracture Mechanics*, Vol. 53, No. 5, pp. 775–788, 1996.

Zhao, Z., Haldar, A. & Breen, F.L., Fatigue Reliability Updating Through Inspections for Bridges. *Journal of the Structural Engineering, ASCE*, Vol. 20, No. 5, pp. 1624–1642, 1994.

Chapter 18

Construction remote sensing, building blackbox, and 3-D laser scanning applications

Liang Y. Liu

Department of Civil and Environmental Engineering,
University of Illinois at Urbana-Champaign, Urbana, IL, USA

ABSTRACT: Advances in information technologies have enabled new approaches to collecting and managing data throughout the life cycle of civil infrastructure systems, such as bridges, highways, mission-critical buildings, and tunnels. Inspections are often conducted to monitor the health of these structures to ensure the safety of residents or general public who utilize the systems. Inspections often involve both intrusive and non-intrusive procedures to evaluate conditions and maintenance needs. Due to budget and resource constraints, most municipal, state, or federal agencies can only conduct periodic inspections subject to available personnel, resources, and budget. As a result, many problems that could have been detected and repaired cost effectively end up as major costly rehabilitation. New technological advances, especially in sensors, mobile computing, 3-D laser scanning, and wireless communications, provide a unique opportunity to rethink the paradigm of infrastructure inspections and how to collect more accurate data for decision making. This paper presents three research projects that utilize these new technologies. The first project explores the use of mobile and wireless technologies for construction inspection tasks, the second project utilizes 3-D laser scanning technologies to accurately document construction progress and as-built data, and the third project develops a building blackbox system to monitor a structure and support emergency response during manmade or natural disasters.

1 Introduction

New technological developments in smart sensors, mobile computing, wireless communications, and 3-D laser scanning have created new opportunities for infrastructure inspections and field data collection. These new developments empower engineers and decision-makers to better analyze the condition of infrastructure systems, such as bridges, tunnels, and underground utilities, with more accurate, long-term, and even real-time data.

This paper highlights these technologies based on three research projects that utilize these advances. The paper is comprised of four main sections. It starts with an overview of the existing research, followed by a discussion of the new enabling technologies. Three recently completed and on-going research projects conducted by researchers at University of Illinois are presented as case studies to highlight these technological advances and their potential applications for infrastructure management.

2 Background and existing research

Since the 1960s, engineers have been using sensors to test or monitor structures such as bridges, tunnels, and dams. Structural engineers have been using sensors to monitor stress, deformation, and dynamic responses under loading conditions. Many structures are outfitted with sensors to monitor their conditions (Tobia and Foutch 1997) or dynamic behaviors (Ko 1996). Schulz (2001) outfitted the Sylvan Bridge in Oregon to monitor deterioration patterns. Maalej (1999) utilized three spectrograms to estimate and compare the fundamental natural frequencies of vibrating structural members. Watkins' (2001) team built a 9.1-meter-long bridge with embedded sensors to investigate its material properties and dynamic behaviors. Success from these efforts has shown the promise in placing sensors as a permanent part of structures. Coupled with the latest advances in mobile/wearable computers and wireless communications, sensor technologies may dramatically change the way we inspect infrastructure systems in the future. The following section summarizes a few of the advances that may change the paradigm of infrastructure inspections.

2.1 *Technological advances available for infrastructure inspections*

Three areas of technological advances are gradually shaping the inspections of the future: sensors, wireless communications, and wearable computers.

2.1.1 *MEMS, fiber optics, acoustic, and chemical sensors*

Fiber optic, acoustic, and chemical sensors, as well as MEMS (micro-electromechanical systems), have provided a variety of ways for detecting the physical and dynamic properties of structural members. MEMS integrate mechanical elements, sensors, actuators, and electronics by way of micro-fabrication technology. These advanced sensing systems operate by transforming the electronic information sent from the sensors to an actuating system that responds to stimulations. These "systems-on-a-chip" enable the development of smart products composed of micro-sensor and micro-actuator systems for detecting the deformation, strain, and stress of structural components. New generations of fiber optic sensors, such as SOFO (Surveillance d'Ouvrage par Fibres Optics-monitoring of structures by Optical Fibres), can detect displacements by utilizing low-coherence interferometry to measure the differences in wave length produced by two optical fibers.

Acoustical sensors can detect sound responses from an excitation on an existing structure. Structural engineers can use the response patterns to correlate strengths or to identify problem areas. These sensors can also detect acoustic emissions from materials under loads/distress heavy enough to cause deformations or cracks.

Chemical sensors send a signal when they detect a chemical reaction. For example, a Schottky diode can be used as a chemical sensor to capture energized electrons and produce a measurable electrical signal. The Schottky sensor is a thin metal film, made of silver, gold, platinum, or another metal sprayed onto a silicon wafer. Scientists have also looked into combining optical sensors with the ability to measure chemical changes. These sensors may be used to detect the existence of ions released by deterioration such as rebar corrosion inside concrete.

2.1.2 *Wireless communications*

Wireless communication has improved significantly in the past ten years, especially local wireless networking (IEEE802.11a/b/g/n and Bluetooth) and satellite communications. IEEE (Institute of Electrical and Electronics Engineers) members and researchers have developed standard protocols for wireless communication networks, which have made it possible for computing devices to communicate with one another. Many mobile and data logging devices are now designed to use these standard protocols so that they can exchange data easily. Currently IEEE802.11a/b/g/n are the most widely used standards for wireless communication networks. Bluetooth is another wireless standard that uses different frequencies and bandwidth. Typically used for short ranges of 10 to 100 meters, Bluetooth has the ability to connect sensor clusters and interface with data loggers to collect inspection data. Satellite communications, such as Hughes Satellite Networks, which utilize VSAT (very small aperture terminal) low-orbit satellites for telecommunications, allow data to be transmitted from the field to remote locations.

2.1.3 *Mobile and wearable computers*

Mobile and wearable computers provide a convenient tool for inspectors to store field data and communicate with remote experts. Commercially available wearable computers, such as Xybernaut Poma and i-Paq from Hewlett Packard, and even iPhones from Apple, make it possible to store digital data in real time directly to local and remote database servers. Figure 1 below illustrates some of the mobile and wearable devices.

Equipped with wireless communication hardware and software, these mobile devices are becoming effective tools to communicate and collaborate with remote experts.

2.1.4 *Mobile and sensor-based infrastructure inspections*

Taking advantage of the aforementioned technologies, researchers at the University of Illinois at Urbana-Champaign are working on a new approach to inspect infrastructure systems such as bridges, tunnels, locks, and dams. The following sections describe the research goal, the schematic design, and preliminary results.

The goal of the project is to utilize new technologies to provide remote and real-time monitoring of civil infrastructure systems. This project explores sensor technologies, including optic fibers, MEMS, and chemical and acoustic sensors for civil engineering

Figure 1 Wearable and hand-held Computers.

applications. Equipped with wireless capabilities, these sensors can communicate with hand-held computers used by inspectors, or directly transmit data over wired/wireless networks and the Internet. The design of the monitoring system provides for both tele-observations and tele-inspections, which support remote access to data in the field and data from the sensors embedded in bridge girders, tunnel concrete lining, and underground foundations. This design enables real-time and long-term monitoring of infrastructure systems, in an accurate and cost-effective manner.

2.1.5 *Sensor-based infrastructure monitoring*

Sensors mounted on the components of infrastructure systems collect and transmit data to remote decision makers via wired and wireless networks and the Internet. Fig. 2 shows the common configuration that highlights some of the new technologies that can be deployed to support infrastructure inspections and long-term data collection. Via local wireless communication networks, data are transmitted from the sensors (accelerometers, strain gauges, and weather monitoring devices) to digital data loggers and from the digital data logger to a remote database server. In remote locations, a two-way VSAT satellite can be used to transmit data between the field and remote server. With Internet connections, remote users can (1) directly interact with inspectors in the field, (2) view and download data from the sensors, and (3) remotely control the

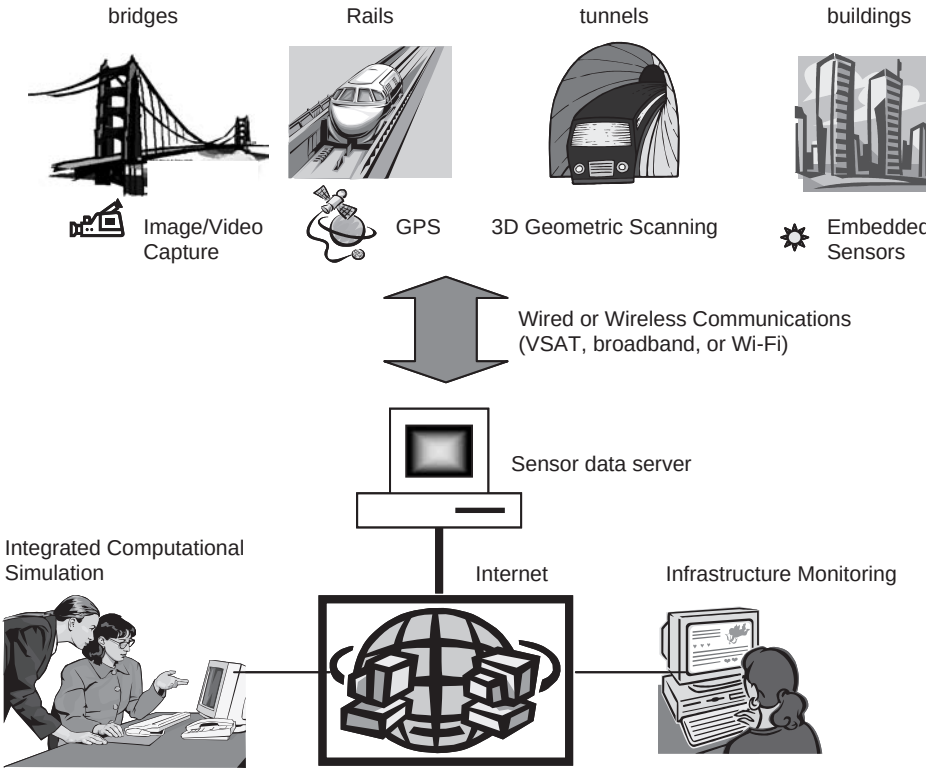


Figure 2 Sensor-based infrastructure monitoring.

view/pan/tilt/zoom functions on the digital cameras in the field. This design optimizes cost, mobility, and future technology upgrades. Each component of the design, such as sensors, communication networks, and database servers, can be upgraded or changed without affecting the overall system functions. This is accomplished by using standard communication protocols and TCPIP (telecommunication protocol and internet protocol).

2.2 *Potential and limitations of mobile and sensor-based technology*

From the aforementioned project, the research team made the following observations on mobile and sensor technologies:

(a) Bright and promising future. We are convinced that sensor-based inspections will greatly reduce the cost associated with infrastructure inspections. The research progress to date convinced us that technologies will change how we design and monitor infrastructure systems in the future. There is a bright and promising future in integrating concrete and steel with sensors and microchips.

(b) Messy wires, wireless interferences, and difficult calibrations. We tested both wired and wireless communications to get data from sensors. In the field, wired connections, which require cables, ports, and connections to data loggers, turned out to be troublesome. Wireless connections are more convenient to deploy; however, unsteady power supplies and interference posed difficulties for our field test. Further, many sensors needed to be calibrated before data could be trusted. The calibration process was lengthy and required experienced personnel. We had to discard a large amount of data due to minor errors in calibration.

(c) Software driver conflicts and hardware compatibility problems. We encountered numerous software driver conflicts when connecting test devices and sensors. Even hardware components purchased from the same vendor were incompatible. This may seem like the quality control problem of one supplier; however, the research team constantly needed to configure hardware and software components from four or five companies. This difficulty proved to be the most problematic and time-consuming to solve.

(d) Fragile sensors and non-rugged instruments. Many existing sensors and instruments for testing structural members are designed for laboratory use. The design assumptions of a dry, clean environment and universal power access, for most sensors and lab instruments, have been a challenge to technicians when we conducted tests in the field. Not surprisingly, many instruments failed and had to be sent for repair, causing many delays to the projects. We also experienced several weather delays and weather-related damage.

(e) Field power supplies and sensitivity. Ensuring stable and reliable power sources turned out to require more than just carrying a generator. We tried both the local power source and power from a portable diesel generator. However, while the subtle fluctuation of voltages and currents might fall within the tolerance of most electronic devices, they caused significant errors on data streams from the sensors. This is because sensors rely on registering changes in voltages and currents to detect properties such as strain, stress, or acceleration.

(f) Smart sensors, wireless communications, and wearable computers will no doubt play a key role in inspections in the future. Sensors–smarter, smaller, and cheaper–will be an integral part of future infrastructures with solar or passive (induced) power. This means in the near future civil engineers will be working closely with electrical engineers and computer scientists in designing structures and infrastructure systems.

3 Laser Scanning Applications in Construction

3.1 *3-D Laser scanning technology*

Three-dimensional laser scanning is a relatively new technology that utilizes LIDAR (Light Detection and Ranging). It is similar to RADAR (Radio Detection and Ranging), but uses light to measure range or distance. Based on the information on the size or shape of the object being scanned, distances can be determined. A laser scanner consists of an emitting diode that produces a light source at a very specific frequency. A mirror directs the laser beam (with a diameter of 6mm) horizontally and vertically towards the target. The surface of the target then reflects the laser beam. Using the principles of pulse time of flight the distance can be determined by the transit time, with a precision of $\pm$ 6 mm. The result of a scan produces point clouds, which can be processed into accurate 3-D models (Kern 2002). LIDAR allows measurement of a large number of points in a relatively short amount of time; these points with accurate dimensions allow engineers to represent excavation site geometry in three dimensions. Fig. 3 shows the scanner and a scanned image as compared to a digital photograph (left).

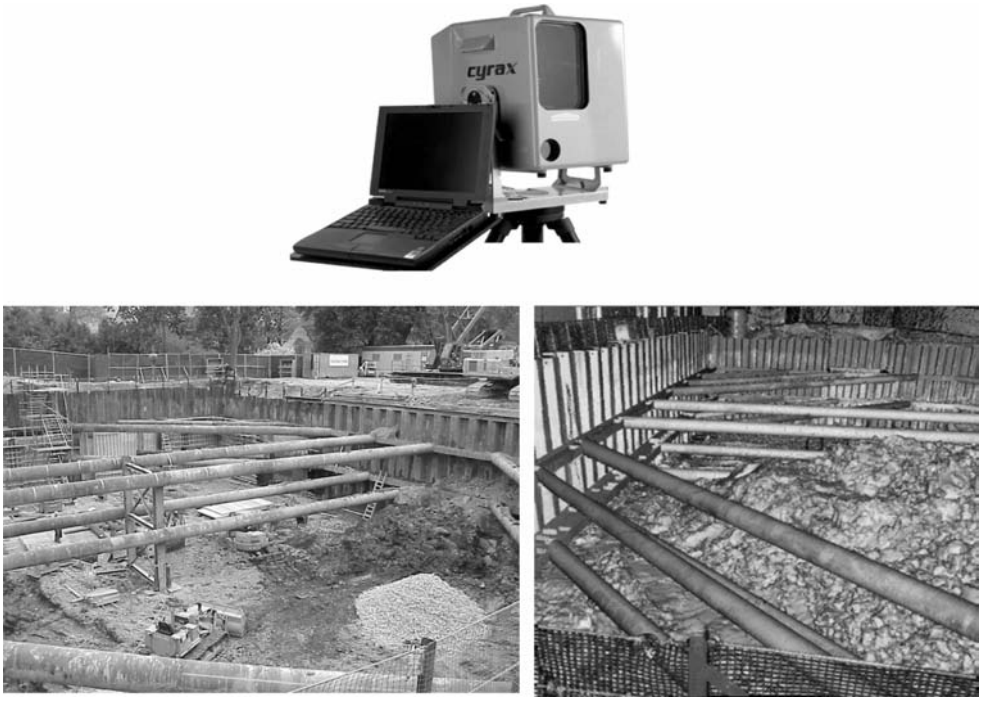


Figure 3 3-D Laser Scanning Technology.

3.2 *3-D Laser scanning application in urban excavation and monitoring*

Urban construction excavations affect the lives of millions of people daily. To minimize the impact on nearby structures, geotechnical and construction engineers must work closely with contractors to design and monitor ground movements. Traditionally, engineers can estimate ground movements using a combination of semi-empirical methods based in part on past performance data and results of model simulation using finite element analyses. However, predictions from model simulations contain uncertainties related to soil properties, support system details and construction procedures. Given these uncertainties, it is common for urban construction sites to install monitoring systems, which record ground movements during construction and, in some cases, movements of adjacent buildings. Significant developments have been made in monitoring systems; examples are automated and wireless systems for measuring lateral deformations using in-place inclinometers and building deformations using automated total stations. This information is useful for geotechnical engineers if it can be matched accurately with accurate construction sequences and activities. However, engineers and contractors continue to rely on limited records of construction field information that are typically collected by hand. There is an obvious mismatch between the quality of data from automated instruments and the record of construction activity.

The following sections describe a study utilizing 3-D laser scanning to monitor construction progress such as pay items, quality controls, and as-built dimensions. The last few sections present lessons learned from the study and provides a discussion for future vision on construction field data collection and engineering integration.

3.3 *3-D In-situ construction data in geotechnical analyses*

Geotechnical engineers rely on in-situ construction data to update and predict ground movements for urban excavations. Using numerical modeling and simulation, geotechnical engineers can analyze and predict static and dynamic soil-structure interaction problems, such as embankment loading, deep excavation, and tunnels. These analyses rely on field measurements of as-built construction sequence, excavation volume, depths, case histories, and soil behaviors. These field measurements provide input to geotechnical analyses that utilize neural network (NN) and finite element analyses (FEM) to understand the dynamics of stresses, strains, and movements. As an example, Figure 4 depicts a process of using field data to predict ground movements during urban supported excavation using the so-called "Autoprogressive" method (Hashash et al 2002). The importance of accuracy of field measurements cannot be over emphasized. Without accurate field data, many predictive models will fail.

In reality, however, collecting accurate field data during construction has not been easy. Equipment movements, personnel training, weather conditions, site characteristics, and other interferences all contribute to difficulties in collecting quality field data. With the advances in information technologies, many new sensors and sensing equipment are becoming available to improve the process of construction field data collection. The following sections will focus on the use of 3-D laser scanning for data collection to support geotechnical analyses.

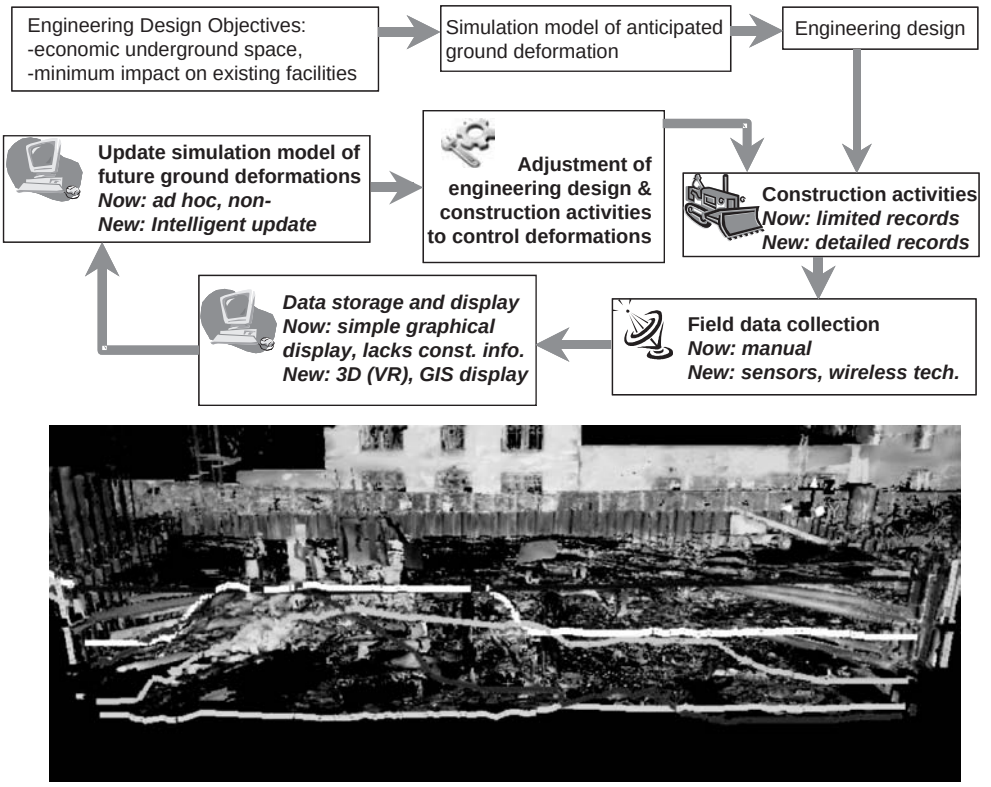


Figure 4 3-D Construction field data for geotechnical analysis.

3.4 *Integration of construction field data and geotechnical analyses*

It is a common practice for urban excavation sites to include a monitoring program during construction to record the ground movements and, in some cases, adjacent building movements. These observations can be used to evaluate how well the actual construction process is proceeding in relation to the predicted movements. Ideally, these observations can also be used to control the construction process and update predictions of movements given the measured deformations at early stages of construction. The weakest link, however, lies in the difficulties in collecting accurate construction as-built data in terms of construction sequence, excavation profiles, and ground support systems. Using an excavation project in the Chicago area, field tests were conducted to verify the accuracy and feasibility of utilizing 3-D laser scanning to collect more accurate construction as-built data for excavation projects. With the scanned images, actual dimensions can be obtained from the 3-D scanned models. As shown in Fig. 4, excavation profile and measurements and terrain model can be constructed and analyzed using finite element analysis software. The 3-D data can also be cross-referenced with other ground monitoring sensors to understand the impact from construction activities.

This integration allows geotechnical engineers to better understand and predict ground movements near the construction sites thereby minimizing the potential impact, delay, and costs of urban excavation projects.

3.5 *Lessons learned*

Several valuable lessons were learned from the research when the 3-D lasing scanning technology was put to the test on a real world project. The following summarizes strengths, weaknesses and opportunities of applying 3-D laser scanning to both construction and geotechnical fields.

Strengths:

- Accurate (± ¼", 8 mm precision) collection of 3-D as-built dimensions of site and as-built components
- Safe and fast data collection process
- Data export capabilities to other A/E/C systems (such as CAD and VRML)

Weaknesses:

- Coordinate system integration among different scan sessions
- Overlapping scans and target selection in progressive scans
- High cost of equipment
- Training needed to become proficient in data manipulation
- High-end computer hardware and software

Opportunities:

- 3-D Model generation and integration with GIS
- Accurate as-built for progress payment in construction claims and disputes
- Automated and remote data capture via the Internet
- Wireless technology and GPS (global position system)

3.6 *Potential research in construction field data collection and automation*

Building upon the lessons learned above, the researchers at the University of Illinois have the following future plans.

3.6.1 *An integrated model for excavation controls and prediction*

From the lessons learned from the field tests, we believe the technologies, despite some weaknesses, will enable engineers and contractors to work closely together to not only enhance the quality of constructed facilities but also minimize the impact on neighboring structures. We envision the engineering and design objectives, such as costs and impact on existing facilities, will be better served by having better models to simulate and predict ground deformations. This prediction will be enhanced by data from construction activities. Instead of analyzing the model with only limited records of construction sequence and as-built data, detailed records will give engineers new ways

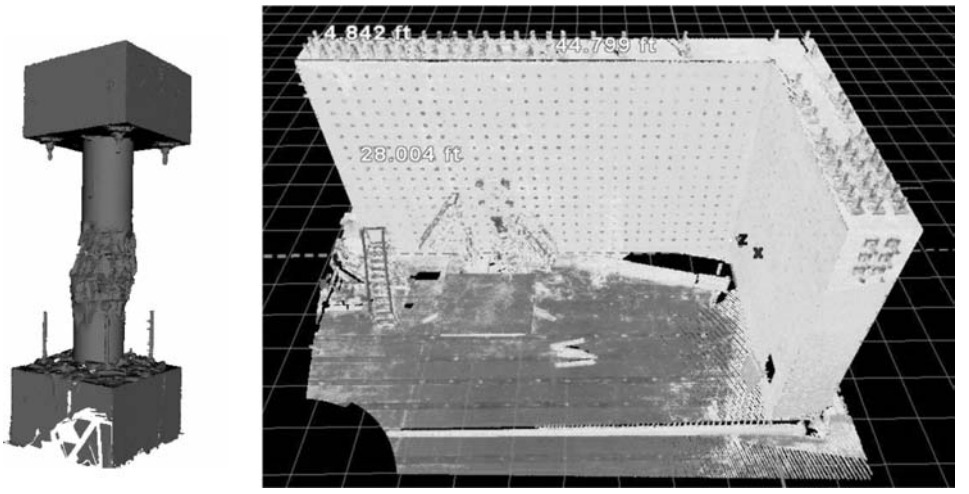


Figure 5 3-D Data for structural failure analysis.

to verify their models and better predict the impact of construction activities. We also envision field data collection to become more automatic and less dependent on manual processes by using sensors, sensing devices, and wireless communications. We expect an integrated GIS to support 3-D virtual reality visualization and simulation. This new paradigm will change the current ad-hoc update of field data into an intelligent and fully integrated environment for predicting and controlling of ground movements.

3.6.2 *Construction automation in inspections, repairs, and pay items*

We also envision that sensors and sensing technologies will continue to change how construction field data are collected. Several promising research areas that we are working on include automatic dimensioning and inspections, repair analyses, and as-built data collection. Fig. 5 shows a scanned wall segment (right) for automatic inspections of dimensions and rebar spacing. A column under repair (left) is scanned for fractures and damages before deigning repair methods. We are also exploring the use of 3-D laser scanning to support other construction cost controls, such as automatic pay item verification.

The research presented in the case study shows the promise of using 3-D laser scanning technology in monitoring urban supported excavation. Using the 3-D laser scanner, geotechnical engineers and construction managers can potentially improve the field data collection process for supported excavation in the urban areas. We believe that acquiring accurate records of construction staging in deep excavations is critical in understanding and predicting the behavior of field excavation as well as impact on nearby facilities. We expect the potential of these technologies will be further realized through the integration of the collected data into engineering analysis processes. 3-D scanning can provide timely as-built data. The terrain model can then be automatically imported into a numerical simulation environment to improve the accuracy of model

simulation. We believe that the integration of accurate field data with analytical models will help improve the quality of design and construction. Although there are still barriers, both technical and non-technical, to overcome, the future is bright for such integration and its potential impact on the industry.

4 Building blackbox system

4.1 *Background*

When natural and man-made disasters strike an urban area, the accurate information regarding the critical physical infrastructures such as building documents and building conditions is highly valuable and critical to disaster relief operations and decision-making processes. Issues related to sharing, searching and routing of such information are critical in enabling the first responders and civil engineers to be aware of rapidly changing conditions in order to develop constructive strategies of action in a disaster environment. The following sections describe the conceptual design of a disaster-survivable building blackbox system that incorporates building sensing and control systems to provide real-time building health conditions. In addition, the proposed blackbox system is equipped with a digital building database which maintains the building documents, such as floor plans and emergency exit routes, to provide on-site preservation of critical information. The blackbox system aims at building an infrastructure of two-way communication between the building systems and the emergency response team over a mobile ad hoc network (MANET) (Aldunate 2006; Frankel 2006; Conti 2005). A field test at the Illinois Fire Service Institute (IFSI) on the preliminary design and prototype implementation of the proposed system showed that building information and documents stored in building blackbox systems were highly accessible and efficient in providing building information to first responders during disastrous events.

4.2 *Impact of disasters and challenges of rescue efforts*

In the past years, several large-scale natural and human-made urban disasters have resulted in tremendous catastrophes and enormous casualties. For example, between 1995 and 2006, a total of 668 disasters occurred world wide affecting over 230 million people, of which 122,260 people died. The average cost of damage is $108 million per disaster (NIST 2005). One of the main causes of these casualties in urban disasters is building collapse or structural damage, for example, the terrorist attack at the World Trade Center complex and the bombing attack at the Murrah Federal Building. Hence, building structure damage and collapse has become one of the most important and challenging issues in Urban Search and Rescue (US&R) operations during disasters involving critical physical infrastructures. However, as some studies indicated, the complex internal structure of the buildings and lack of real-time, accurate and reliable building information available to both emergency responders and people inside the buildings made it particularly difficult for either rescue operations or quick evacuation in an emergency situation (NIST 2005; Kwan et al. 2005). Due to their expertise, civil engineers and constructors are being considered as the "fourth responders" to provide the rescue team with critical information to support resource allocation, risk assessment and decision-making. Nevertheless, critical building information, such as

building design, potential building failure and victim locations, is often missing and difficult to obtain, hampering rescue efforts. As recommended in the Final Report of the National Construction Safety Team, a system for "*real-time secure transmission of valuable information from fire alarm and other monitored building systems for use by emergency responders, at any location, to enhance situational awareness and response decisions and maintain safe and efficient operations*" should be developed (NIST 2005).

To support and improve the disaster relief efforts, we propose a building blackbox system which is designed to be disaster-survivable and highly accessible by utilizing state-of-the-art high temperature and high strength geopolymer material as well as insulation technology. Equipped with a building information database that employs certain degrees of data redundancy, this building blackbox system aims at providing critical building information and facilitating emergency response to disasters on complex building structures in an efficient and effective manner (Fig. 6).

4.3 *Building blackbox system critical information acquisition*

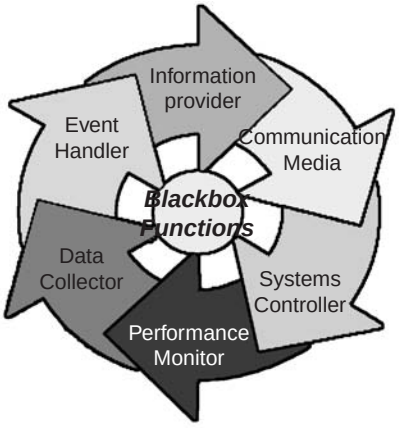
To make the rescue operations in large-scale building emergencies more effective and efficient, the proposed building blackbox system is designed to be integrated with building sensing and control systems to provide critical and valuable information in a timely manner regarding the status of critical physical infrastructure including

(a) internal structural health of the building (e.g. stress and strain data of structural elements);
(b) disaster environment conditions and building performance (e.g. audio/video/image data, temperature, humidity and air quality);
(c) building systems status (e.g. statuses of communications and HVAC systems);
(d) personnel (both occupants and first responders) location and health information, which could be provided by a wearable physiologic monitoring system (Kwan et al. 2005).

By gathering, processing and delivering the above-mentioned critical information to first responders, the building blackbox system acts as an information hub that will be particularly useful to enhance the situational awareness of all emergency responders (NIST 2005).

4.4 *Building information database*

In addition to the data from sensor and control systems, the building blackbox system also has a building design and construction database designed to be easily accessible by emergency responders during disaster events. In this database, construction drawings of floor plans and layout, completed structure material specifications (bill of material reports), deployment of electrical, plumbing and fire protection systems, as well as drawings of emergency exit routes are compiled and indexed for quick search queries (Prieto 2002). This information plays an important role in the evacuation process, the rescue planning, and the engineering assessments.



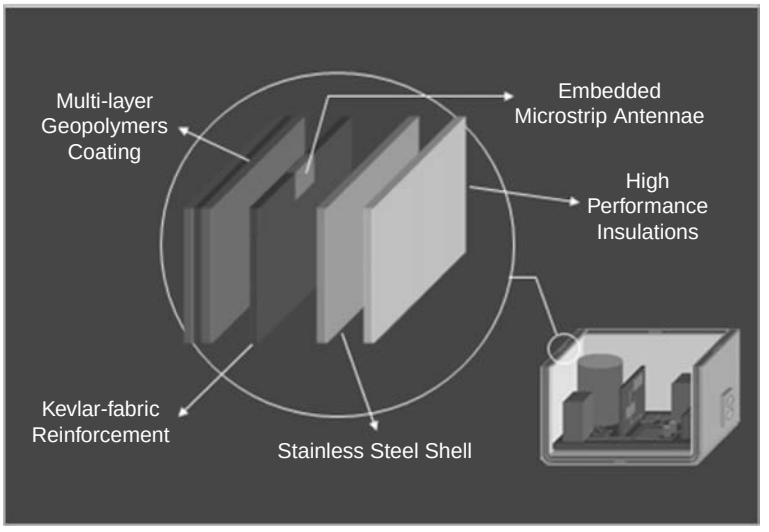


Figure 6 Blackbox system functions and design.

4.5 *Disaster-survivability*

Designed to be a critical information provider in disaster situations, the building blackbox system must survive various disaster scenarios. Fires, static crushes, crash shocks, vibrations, fluid immersion and any combination of these events are likely to happen to a building and the blackbox systems in natural, accidental or intentional events. Table 1 depicts the design specifications of the survivability requirements for the blackbox systems. These requirements are derived according to NIST standards (NIST 2005). However, some requirements were derived from other sources such as FAA aircraft recorder specifications and rail and locomotive recorder specifications

Table 1 Survivability requirement for the building blackbox system.

Situation	*Typical environment*
Fire (High temperature)	1200°C (RC Structure) for 60 minutes
Fire (Low temperature)	260°C for 10 hours
Static Crush Pressure	5000 lbf for 5 minutes (faces and diagonals)
Shock	23 g's/250 ms or energy equivalent
Penetration	500 lbs with 0.05 in^2 steel pin drop from 10 ft height
Immersion (Fire extinguisher fluid)	8 Hours
Immersion (Corrosive fluids)	48 Hours

(Thompson et al 1999) in the absence of definitive information. This set of specifications are set as the starting criteria for designing the building blackbox system.

Among these requirements, emphasis was placed on high temperature and static crush pressure requirements in the preliminary design since they are mostly likely to be encountered in building disasters. To ensure the survivability of the building blackbox system for these requirements, three state-of-the-art technologies including high temperature and high strength geopolymer material, insulation technology and the solid-state memory unit will be explored and discussed in the following sections.

4.6 *Geopolymer protection*

Geopolymer, or "man-made rock", is a noncombustible, heat/fire/acid resistant material synthesized by raw materials (e.g. fly ash), inactive filler (e.g. metakaolinite) and geopolymer liquor (e.g. potassium silicate solution). Due to its excellent durability performance and high bond strength with steel, concrete and wood, geopolymer has been widely used as a protective coating material (Balaguru 1998). In our experiments, Kevlar-fabric reinforced geopolymer composites have two excellent properties: high thermal stability (up to 1528°C) and high compressive strength (up to 100 MPa) (Duxson et al. 2005). These two rugged properties meet the survivability requirements for the blackbox system. In addition, another critical advantage of using geopolymer composites as the coating material for the building blackbox is that it could avoid shielding problems of wireless signals which metal housing might encounter. These properties make the Kevlar-fabric reinforced geopolymer composites excellent material for the ruggedized protection in building disaster environments.

4.7 *Insulation*

To maintain the interior environment of the building blackbox system at a workable temperature and to prevent the damage of electronic components from overheating, the disaster-survivable design of the building blackbox system is internally insulated with dry-silica based thermal insulation. Like those in the aircraft blackbox and space shuttle thermal protection system (TPS), the high performance insulation material to be used in the building blackbox system has a service temperature up to 1260° C while

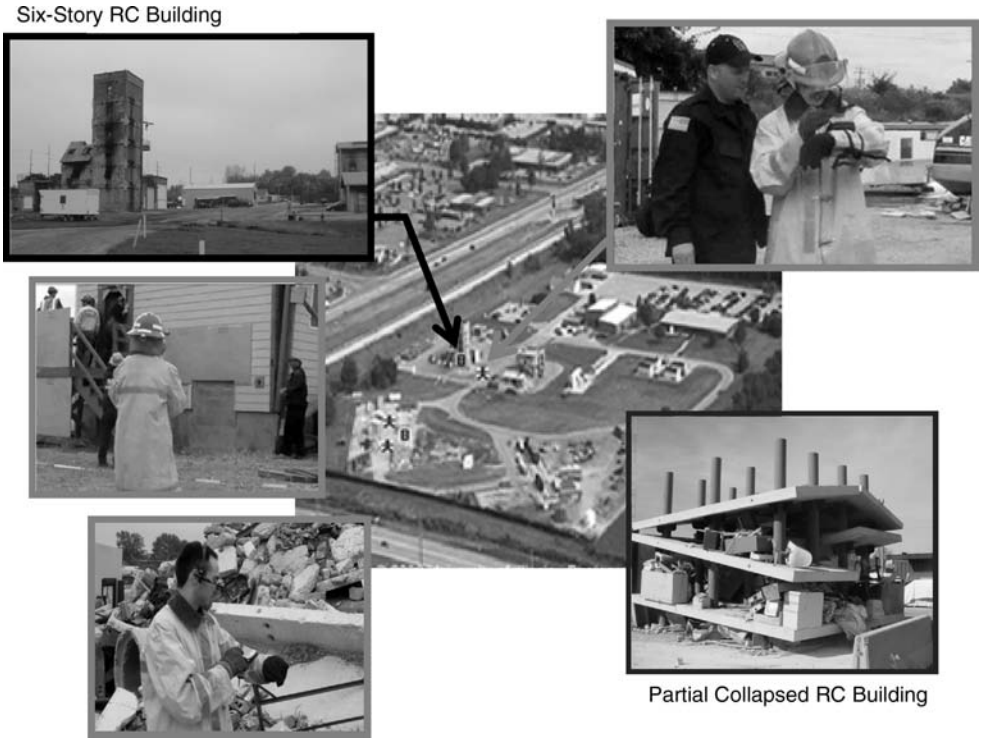


Figure 7 Blackbox system field tests.

its low thermal conductivity would reduce flow of heat transmitted into the interior to a tolerable amount.

4.8 *Solid-state memory unit*

Another state-of-the-art technology used as part of the disaster-survivable design is the solid-state memory unit. A solid-state memory unit is a highly reliable flash-memory based replacement for conventional hard-disk drives; it has very complex modules of internal NAND non-volatile memory chips and controllers instead of magnetic, spinning disks.

4.9 *Field tests*

To evaluate the efficiency and the effectiveness of the proposed system, a field test on a prototype was carried out at the Illinois Fire Service Institute (IFSI). In a series of field test, the research team verifies and validates the concepts developed for the on-going research of the blackbox system. Figure 7 shows the field test configuration with researchers working closely with firefighters and accessing information such as structural analysis, building electrical drawings, maintenance records, and temperature data from the sensors (Figure 8). These field tests further confirm the importance and

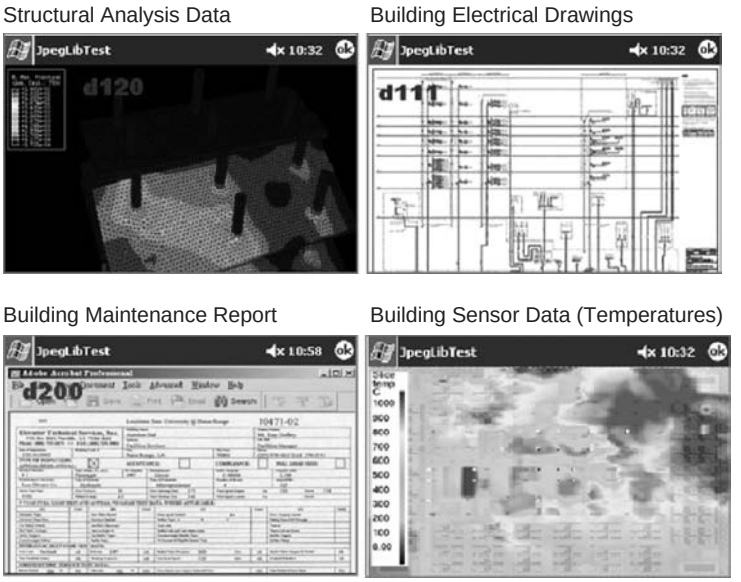


Figure 8 Blackbox system data access.

viability of the blackbox system. The researchers use these field tests to further refine the design of the blackbox system.

5 Infrastructure life-cycle data for facility management and risk assessment

With the aforementioned sensor technologies, decision makers for infrastructure management and risk assessment can have access to more accurate and even real-time information. The ability of these advanced technologies to provide long-term accurate information will enable policy makers to better utilize limited resources for maintenance. Designers can also rely on the data to verify structure performance. With better data, risk assessment becomes more transparent because of the availability of comprehensive and objective data. It will be commonplace in the future to rely on data streaming from the field in real-time to make maintenance decisions and to remotely trouble-shoot infrastructure problems from autonomous diagnostic sensor systems embedded in critical components of civil infrastructure systems. Throughout the life cycle of design, construction operation and maintenance, sensors will need to be designed, maintained, and replaced (Fig. 9). Civil engineers will work closely with other engineering disciplines such as electrical engineering and computer science to jointly define the functionality, design and install the sensors, in addition to traditional design and construction of concrete slabs and steel girders. The long-term accumulation of data/information/knowledge will lead to more innovation in engineering design. Furthermore, long-term data from sensors can greatly enhance the accuracy of probabilistic assessment (Ang and Tang 2007) of infrastructure system performance and structural reliability under seismic conditions (Wen and Song 2003).

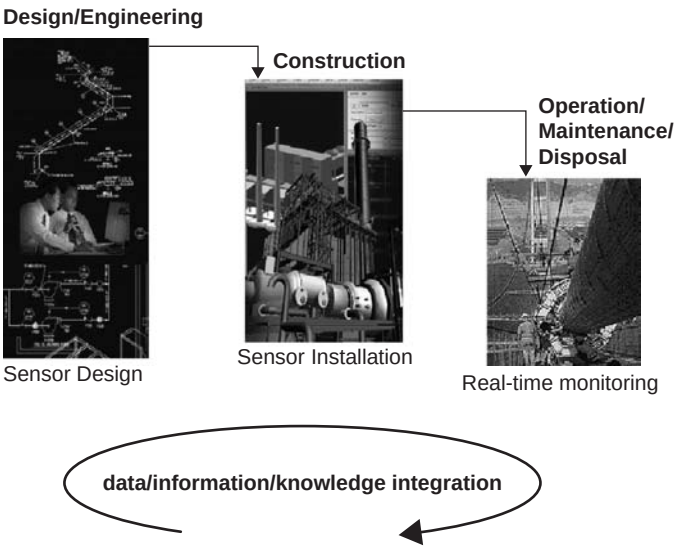


Figure 9 Infrastructure life-cycle integration.

6 Conclusion

Information technology tools, such as sensors, 3-D laser scanning, wireless communications, and mobile computing, can and will become an indispensable part of construction field data collection. These tools can reduce time and enhance quality in collecting field data. Undeniably, new technologies come with uncertainties, risks, costs, problems, and resistance. Companies interested in implementing IT within their organization must recognize the nature of new technologies, their life cycle, and most importantly how to integrate work processes with technologies. Adopting technologies for technology's sake often leads to bleeding edges and failures. Not all new technologies survive the true test of the real world. It is imperative that the industry, technology developers, and researchers work closely together so that basic research will lead to applied fields that will benefit the industry directly, shortening the time for technology transfer. Infrastructure management and risk analysis both require quality data to support decision making. Sensors, 3-D laser scanning, and mobile computing all offer the potential of collecting real-time and long-term data directly from the field. These technologies will continue to play a pivotal role in the future of infrastructure management. However, to implement these technologies cost-effectively, one must remember to fully understand the strengths and weaknesses of each technology and to use small-scale field tests to iron out potential problems in large-scale implementation. The future could not be brighter if the researchers, technology developers, and the industry could collaborate to reduce risks, uncertainties, and costs of technology development, transfer, and usage. As a last note, we need to keep in mind that education will be critical to train the next generation engineers and infrastructure managers with the knowledge not only in civil engineering fields but also basic knowledge in sensors so that they can collaborate with other engineering disciplines that will work

closely with civil engineers in designing and constructing infrastructure systems in the future.

Acknowledgement

The research presented in this paper is the result of dedicated efforts by many graduate students whom the author had the privilege to advise and work with. They include Sangyoon Chin, Chris Erickson, Jeff Monk, Chulsoo Kim, Katrina Ballado, Torsten Trupp, Kevin Foster, Vance Tsai, Ricky Su, and Seokyon Hwang. The materials for the paper are based upon works supported by the National Science Foundation CMS (Civil and Mechanical Systems) Division and ITR (Information Technology Research) Initiatives. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author's and do not necessarily reflect the views of the National Science Foundation.

References

Ang, A. H-S, and Tang, W.H., *Probability Concepts in Engineering*, 2nd Edition, J. Wiley & Sons, Inc., New York, 2007

Aldunate, R., Ochoa, S. F., Pena-Mora, F., and Nussbaum, M.. Robust Mobile Ad Hoc Space for Collaboration to Support Disaster Relief Efforts Involving Critical Physical Infrastructure, *Journal of Computing in Civil Engineering*, vol. 20, pp. 13–27, 2006

Balaguru, P. Geopolymer for Protective Coating of Transportation Infrastructures, *Technical Report, CAIT/Rutgers, the* State *University*, New Jersey, 1998

Conti, M., Gregori, E., and Turi, G. A Cross-Layer Optimization of Gnutella for Mobile Ad hoc Networks, *Proceedings, 6th ACM MobiHoc Symposium*, Urbana-Champaign, Illinois, USA, May 2005, pp. 343–354, 2005

Duxson, P., Provis, J. L., Lukey, G. C., Mallicoat, S. W., Kriven, W. M., and van Deventer, J. S.J. Understanding the relationship between geopolymer composition, microstructure and mechanical properties, *Colloids and Surfaces A – Physicochemical and Engineering Aspects*, vol. 269, pp. 47–58, 2005

Frankel, S., Eydt, B., Owens, L., and Kent, K. *Guide to IEEE 802.11i: Establishing Robust Security Networks (Draft)*, National Institute of Standards and Technology (NIST), Gaithersburg, MD, 2006

Hashash, Y.M.A., Marulanda, C., Ghaboussi, J. and Jung, S., Systematic update of a deep excavation model using field perfromance data, *Computers and Geotechnics*, 2002.

Kern, F., Precise Determination of Volume with Terestical 3-D-Laserscanner. *Geodesy for Geotechnical and Structural Engineering II,* 531–534, 2002

Ko, Wen K., The future of sensor and actuator systems." *Sensors and Actuators*, V. 56, 193–197, 1996

Kwan, M. -P., and Lee, J. Emergency response after 9/11: the potential of real-time 3-D GIS for quick emergency response in micro-spatial environments, *Computing, Environment, and Urban Systems*, vol. 29, pp. 93–113, 2005

Maalej, M., A. Karasaridis, D. Hatzinakos, and S.J. Pantazopoulou, Spectral analysis of sensor data in civil engineering structure, *Computers and Structures*, V. 70, 675–689. 1999.

NIST *Federal Building and Fire Safety Investigation of the World Trade Center Disaster: Final Report of the National Construction Safety Team on the Collapses of the World Trade Center Tower*, NIST, U. S., 2005

Prieto, R. Keynote Presentation, FIATECH, *Homeland Security and Roadmap Implementation Workshop*, Nov. 13–15, National Conference Center, Lansdowne, VA, 2002

Schulz, Whitten L., Joel P. Conte, and Eric Udd. Long Gage Fiber Optic Bragg Grating Strain Sensors to Monitor Civil Structures, *Proceeding of SPIE: Smart Structures & Material 2001: Smart Systems for Bridges, Structures, and Highway*, V. 4330, 56–65, 2001

Thomas, S., Onley, R. E., and Morich, R. S. Design of a Crash Survivable Locomotive Event Recorder, *International Symposium on Transportation Recorders*, NTSB, pp. 313–323, 1999

Thompson, M. H. A Vision of Future Crash Survivable Recording Systems, *International Symposium on Transportation Recorders*, NTSB, pp. 337–349, 1999

Tobias, D.H. and Foutch, D.A., Reliability Based Method for Fatigue Evaluation of Railway Bridges, *Journal of Bridge Engineering*, ASCE, V. 2, No. 2, 53–60, 1997

Watkins, Steve E. and Unser, John F. Antonio Nanni, K. Chandrashekhara, and Abdeldjelil Belarbi. Instrumentation and manufacture of a smart composite bridge for short-span applications. *Proceeding of SPIE: Smart Structures & Material 2001: Smart Systems for Bridges, Structures, and Highway*, V. 4330, 147–156, 2001

Wen, Y. K. and Song,S.-H., Structural Reliability/Redundancy under Earthquakes, *Journal of. Structural. Engineering*, Volume 129, Issue 1, pp. 56–67, 2003

Chapter 19

Understanding and minimizing risk associated with dynamic earthwork operations using discrete event simulation

Julio C. Martinez

School of Civil Engineering, Purdue University, West Lafayette, IN, USA

ABSTRACT: An important part of infrastructure engineering involves the design of complex construction operations. Often, such operations involve considerable risk due to inaccurate understanding of the cost and time required to accomplish them. Discrete-Event simulation modeling and visualization can be of significant assistance in understanding the associated risks, and in reducing these risks via improved operations design. This paper illustrates these concepts in the context of earthwork operations. The paper describes the essential characteristics of earthwork operations and the thought processes that guide their design. It also shows how to use discrete-event simulation to design these operations so that involved risks can be better understood and minimized. The operation selected to illustrate the design process includes two interesting aspects: a long and narrow curve that can only accommodate traffic in one direction at a time, and the routing of trucks to one of two possible excavators. Both aspects require the implementation of dynamic strategies for their improvement and present complexities that can be analyzed well using discrete-event simulation. the field applicability of the various strategies explored and the effectiveness of traditional solutions are also discussed.

1 Introduction

Due to the highly dynamic nature of complex earthworks, there is significant uncertainty associated with the time and cost required to complete an operation, unless it is very carefully analyzed using methods that capture the inherent variability of the activities that comprise the operation and the logistics that guide its implementation in the field. Quite often, operations to take much longer and cost much more than anticipated, significantly increasing the risk of completing the work on time and within budget. In order to make appropriate risk aware decisions, it is important to both understand the nature of the risks involved, and to improve the design of the operation so that these risks are minimized.

Discrete-event simulation (DES) is a modeling technique that has been used to analyze and design many construction operations. DES is particularly beneficial for modeling complex dynamic systems that are intractable to other modeling approaches. The capability of DES to address complex dynamic problems, however, is not obvious.

This paper illustrates the use of DES in the design of a complex and dynamic earthwork operation.

2 Earthwork

Earthwork operations involve the excavation, transportation and placement or disposal of materials. They typically involve repetitive work cycles, expensive fleets and large volumes of work. Consequently, even small improvements in planning result in substantial cost and time savings. It is for these reasons that earthwork operations improvement has been the focus of so many studies.

The work is performed outdoors under conditions that are highly variable and that affect the performance of the different pieces of equipment. Factors that affect performance include weather (i.e., trucks cannot travel on wet, muddy haul roads as well as on dry, well graded surfaces), haul road maintenance (i.e., a well-maintained road reduces rolling resistance), operator experience, ground conditions, load and dump area layouts, and the material being excavated. These issues are described in more detail in (Schexnayder et al. 1999) and (Gransberg 1996).

Earthwork operations are actively managed with many decisions made dynamically on site in reaction to the evolving status. A truck, for example, may be routed to an alternate load area if the loading unit at the main area is under maintenance or if several other trucks are queuing for it. Sometimes the strategies that guide these dynamic decisions are quite complex but necessary. Their impact on the performance of the operation may be significant.

The probabilistic nature of the work and the dynamics of earthwork operations make them difficult to plan. In lieu of detailed analysis, they are often planned using simplified back-of-the-envelope calculations, but rely mainly on the experience and insights of the planner. Discrete-event simulation is an earthwork analysis method that can explicitly incorporate the detailed but significant aspects (e.g., equipment characteristics, haul road conditions, load and dump area configuration, and dynamic context-based decisions) of an operation.

3 Earthwork design case study

The case study presented here is a combination of two separate operations that took place in Virginia and North Carolina. These two operations have been combined to make it possible to illustrate two separate but interesting issues of relative complexity in a single exposition.

3.1 *Problem statement*

The operation involves moving 975,000 bank m^3 of 2000 bank kg/m^3 material (1,300,000 loose m^3 of 1500 loose kg/m^3 material) in 75 workdays (16 work hours each) from two possible sources to a common dumpsite as shown on the plan view of Figure 1. The two sources are located towards the bottom left part of Figure 1 and are labeled *MLA* (main loading area) and *ALA* (alternate loading area). The dumpsite is towards the top right part of Figure 1 and is labeled *DumpArea*. The haul distance from the main and alternate loading areas to the dump area are 1,670 meters and 1,920 meters. Both haul routes share 1,370 meters and include a narrow curve 470 meters in length. The narrow portion is not wide enough to allow simultaneous traffic in both directions. Due to the obstruction shown on Figure 1 and other site constraints, it is not feasible to widen the curve.

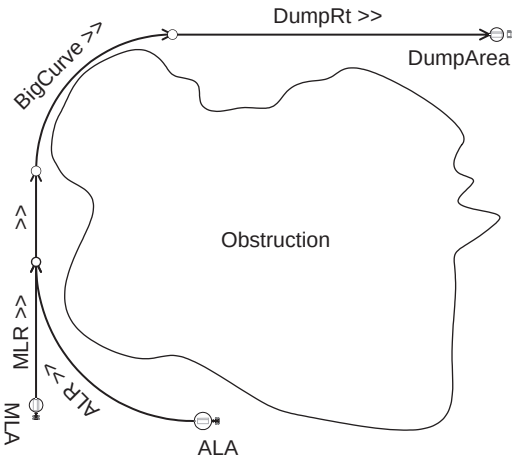


Figure 1 Plan view of earthworks operation site.

Table 1 Haul route segment characteristics.

Segment	*Length (m)*	*Grade %*	*RR %*
MLR	300	5.7	4.0
ALR	550	8.0	5.0
LdRts	200	0.0	3.0
BigCurve	470	3.0	5.0
DumpRt	700	1.6	3.0

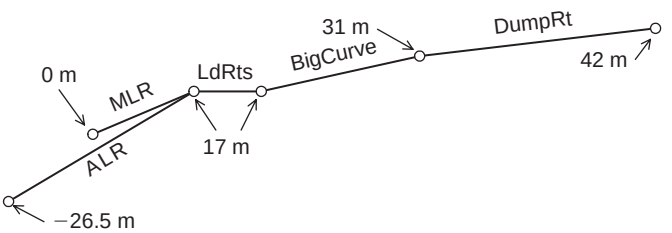


Figure 2 Vertical profile of haul routes.

The dump area is 42 meters above the main loading area and 68.5 meters above the alternate loading area. Figure 2 shows the vertical profile of both haul routes.

Table 1 shows the grades, rolling resistances, and relative elevations of each segment. The underfooting in several parts of the haul routes is soft. The maneuvering space at the load and dump areas is limited.

In order to keep this discussion focused on the issues of dynamic logic, the impact on the operation of machine breakups, repairs, and maintenance are not considered.

3.2 *Preliminary deterministic analysis by hand*

The initial decision is to determine the equipment that will be used for the operation based on their suitability for the task at hand and availability. Poor underfootings demand haul units with high flotation and limited spaces require high maneuverability. Although off-road rigid trucks would be desirable due to the large capacities available, their maneuverability is limited and cannot work well on very soft ground. Articulated trucks, although typically smaller in capacity, are specifically designed for soft ground and high maneuverability, and are thus the clear choice for this operation. In addition, contractors must often choose hauling fleets consisting of equipment they already own. The purchase of new equipment is only justified if their suitability for the task is clearly superior to the equipment already owned, and to the extent that it makes economic sense. The contractor's largest articulated trucks were Caterpillar D30Ds (17.2 m^3 heaped 2:1 SAE capacity). Although slightly larger articulated trucks are available in the market, the D30Ds appear to be a good choice due to their availability.

The best match to the D30Ds from those loading units available to the contractor is a Hitachi EX1100 excavator with a 5.5 m^3 bucket and a 1.065 fill factor (for this material). In three passes, this excavator can load an average of $5.5*1.065*3 = 17.6$ loose m^3 (lcm). The required rate of production for this job is 1,300,000 lcm in 1200 hours, or 1083 lcm/hr. Under the conditions at the main loading site, the best estimate of a swing-excavate-swing-dump cycle for an EX1100 is about 0.41 minutes, yielding a maximum production at 100% utilization of about 857 lcm/hr. A single EX1100 at the main loading site is therefore not capable of sustaining the required production and must be supplemented with another loading unit at the alternate loading area. Because the route to the alternate loading area is longer and steeper, the excavator there should be used only to supplement the EX1100. It is thus convenient to use a smaller unit that can still match the D30Ds, such as the available Caterpillar 375 excavator with a 2.8 m^3 bucket and a 1.06 fill factor (for this material). In six passes, this excavator can load an average of $2.8*1.06*6 = 17.8$ lcm. Under the conditions at the alternate loading site, the best estimate of a swing-excavate-swing-dump cycle for a CAT 375 is about 0.33 minutes, yielding a maximum production at 100% utilization of about 540 lcm/hr. Thus, the combined maximum production capability of the EX1100 and CAT 375 is 1397 lcm/hr; about 29% higher than the minimum required 1083 lcm/hr.

The cycle time for a D30D traveling exclusively between the main load site and the dump area is about 10.0 minutes, assuming constant operation, no waiting times, and a truck dumping time of 0.8 minutes. This yields a production of 17.6 lcm/10.0 min = 106 lcm/hr. When traveling exclusively to the alternate loading area under similar conditions, the cycle time is about 13.3 minutes and the corresponding production is 80 lcm/hr. It is obviously best and most productive to use the main excavator as much as possible, which under ideal conditions can serve up to 8 trucks for a production of 845 lcm/hr. The alternate excavator could supplement by serving 4 trucks for a production of 321 lcm/hr. This gives a total combined production of 1166 lcm/hr, or 8% above the minimum 1083 lcm/hr that are required. Sustaining this level requires about 83% of the maximum capacity provided by the two excavators.

The actual haul road can also limit production. In this study, the narrow curve is the most constrained and its maximum transport capacity needs to be analyzed. Assuming

the direction of travel in the narrow segment changes after every n trucks enter, the capacity of each truck is c, the time to travel the segment loaded is tl, the separation between trucks traveling loaded is sl, the time to travel the segment empty is te, and the separation between trucks traveling empty is se; the production rate (PR) of the segment at maximum utilization can be computed to be:

$$\mathrm{PR} = c^{*}n/[tl + sl^{*}(n-1) + te + se^{*}(n-1)] \quad (1)$$

With $c = 17.7\,\mathrm{m}^3$, $tl = 1.87$ min, $sl = 0.12$ min, $te = 0.71$ min, and $se = 0.05$ min, the production in m^3/hr as function of n is $1062^{*}n/(2.41 + 0.17^{*}n)$. This indicates that the transport capacity of the narrow segment increases with n, resulting in 412 lcm/hr for $n = 1$, 772 lcm/hr for $n = 2$, 1091 lcm/hr for $n = 3$, 1375 lcm/hr for $n = 4$, and 2864 lcm/hr for $n = 12$. Thus, with n set to at least 3, the narrow segment has adequate transport capacity to support this operation.

Although the transport capacity of the narrow curve increases with n, this increased bunching is known to degrade overall system performance. Ideally, the more evenly spaced the haul units, the better the production of the system as a whole.

4 Simulation of the operation

A diverse range of tools can be used to model an operation using discrete-event simulation. Listed in increasing order of flexibility, power, and training required for their use, these tools include: 1) special purpose tools that target a narrow domain such as pipe laying, earth moving, or concrete placement, 2) general-purpose simulation systems that may be geared towards a specific industry, and 3) general-purpose programming languages.

The tool used to simulate the models described in this paper is StroboDirt (Martinez, 1998), which relies on Stroboscope (Martinez 1996) for the actual simulation work. The specific details of this tool are out of the scope of this paper. Regardless of the tool used, the operation must be defined in such a manner that it is communicated unambiguously and completely.

Whenever there is choice in the routing of a truck, it is necessary to define how this choice is made. For trucks coming from *MLR* or *ALR* the choice is fixed such that the next segment traversed is *LdRts*. For trucks coming from *LdRts* a dynamic choice is necessary since sometimes the routing will be to *MLR* and other times to *ALR*. In the initial design of this operation, there is a 66.7% probability that the next segment is *MLR* and a 33.3% probability that the next segment is *ALR* (i.e., randomly route 66.7% to the main excavator and 33.3% to the alternate excavator). This initial truck routing strategy is rather naïve but is the easiest to set up. Subsequent stages of this design, described later in this paper, explore more sophisticated truck routing strategies by using dynamic calculations to define likelihoods.

The loose material density ($1500\,\mathrm{kg/m}^3$ used for both loading areas in this example) is used to calculate the weight of the load carried by the trucks. The fill factor is defined with a probability distribution that enables modeling variations in the amounts excavated relative to the excavating unit's bucket size. Larger values can be used for cohesive materials that fill well and smaller values can be used for dry granular materials that do not. In this operation Uniform[1.05,1.08] (i.e., uniformly distributed

between 1.05 and 1.08) and Uniform[1.05,1.07] are used for the main and alternate excavators. A probability distribution that adjusts the digging time enables modeling variations in the time to excavate relative to the standard digging time supplied in the specifications for the excavating unit. The variance of the distribution reflects the consistency of the operator or the material. The mean of the distribution reflects the ease with which the material can be excavated. In this operation, Triangular[0.8,1,1.4] (i.e., triangularly distributed with minimum 0.8, mode 1, and maximum 1.4) is used for both loading areas. Another probability distribution adjusts the bucket dump time relative to the standard time for the excavator. Large values can be used to model sticky material and low values to model material that flows out of the bucket easily. In this operation, Uniform[0.9,1.1] is used for both loading areas. Probability distributions for truck entry and exit times to the area enable modeling the ease, difficulty and consistency associated with maneuvering. In this operation, Uniform[0.15,0.20] and Uniform[0.08,0.12] minutes are used as entry and exit times for both loading areas. A positioning factor is used to adjust the time for a standard 90-degree swing to the actual conditions. Distributions with means larger than one, for example, can be used for swings larger than 90 degrees or for conditions where there is poor visibility. In this operation, Triangular[0.9,1,1.2] is used for both loading areas.

Characteristics of dump areas include the maximum number of simultaneous dumps and a probability distribution that adjusts dumping time from the standard time associated with the truck. Distributions with high means and variance, for example, can be used to model difficult and inconsistent dumping maneuvers. For this operation the maximum number of simultaneous dumps is limited to one and the dumping time factor is set to Uniform[1.4,1.8].

The fastest time that can be achieved by an experienced operator who operates the truck as aggressively as possible can be determined using engineering principles. This calculated speed can be adjusted by multiplying by a probability distribution associated with each segment in order to model operator inconsistency, experience, and aggressiveness. For this operation, the factor Uniform[0.85,1.0] was used for all segments. The ability or inability of trucks to pass each other is also something that impacts the operation. Typically, passing is not allowed in off-road earthmoving operations under normal circumstances, and the models explored here adhere to that practice.

Since *BigCurve* is a narrow segment that allows travel in only one direction at a time, the rules used to allow/disallow trucks to enter from either side must be defined. The base, naïve, operating logic for the *BigCurve* segment is such that if a truck arrives when the segment is empty, the truck is allowed entry and establishes the current direction. This direction is maintained as long as trucks keep arriving at the same end of the segment. Trucks eventually stop arriving at that end and the curve clears as the last truck exits. At this point direction of travel reverses if trucks arrived and are waiting at the other end. Otherwise it is again established by the next truck to arrive. This operating logic is equivalent to using a very large value of n in equation (1), and is set as default because it yields the highest transport capacity at full utilization (i.e., if it is a bottleneck). As will be shown later in this paper, the entry to the BigCurve segment needs to be defined in a dynamic manner based on current state information in order to implement complex logic. For now, the entry into *BigCurve* remains as described here.

The number of hours for which to run the simulation is typically an important parameter that must be chosen with care. In this example, simulations are set for 16 hours because that is the length of the workday. Although the production of each day is stochastic, that day is assumed typical and the total time of the entire operation can be determined from the 16-hour production rate. Full 75-day simulations are not practical at this stage of the design because several thousand alternatives may still need to be tried. In order to obtain the best possible estimate of true performance with a very limited number of runs, each alternative was run twice using a pair of antithetic runs. The results reported are the average of both runs. In order to make the comparison between alternatives more realistic, they all used the same random number seed and streams (common random numbers). Law and Kelton (2000) describe the techniques of antithetics and common random numbers in detail. Ioannou and Martinez (1996) illustrate the technique of common random numbers in the comparison of alternative methods of tunnel construction.

Other important aspects of the design are the number, type, initial location and direction of the hauling units. This operation starts with twelve D30D articulated dump trucks entering *LdRts* from the junction with *BigCurve*.

Typical simulation output includes tables, charts, and dynamic output in the form of animations. The most important result is the production rate of the operation. Tables and charts describe the distribution of waiting times at different parts of the haul system, the utilization of the loading units, and the distribution of travel times at each segment.

5 Analysis and improvement of the operation

The operation, when simulated as described above, produces the results shown in the top row of Table 2 (production of 764.3 lcm/hr). There is some obvious inefficiency in the operation. The utilization for both excavators is low while the waiting times for trucks entering the curve empty and waiting for the loading units is high. Only 71% of the required production is achieved despite a design that at first impression appeared to be able to produce 8% above requirement. **Obviously, planning on the basis of the non-simulation based analysis would have exposed the contractor to significant risks in being able to complete the entire project on time and within budget.**

A better insight into the operation and its problems are obtained by looking at the graphical output. Perhaps the most interesting is the chart shown in Figure 3. This chart shows the number of trucks waiting to enter the curve empty as a function of time.

Notice that the chart looks like the skyline of a city with tall but separated skyscrapers. Trucks bunch up before entering the curve empty until they are finally allowed to enter one after the other. Charts for trucks waiting at the loading areas would be very similar. Charts for the utilization of the loaders show that they are very active for periods of time (serving several bunched up trucks), followed by periods of inactivity waiting for another bunch of trucks.

An animation of the operation provides a much better picture of the truck bunching and additionally reveals strategies that could be used to improve the operation. Figure 4 shows a snapshot of the animation with 6 empty trucks bunched up waiting to enter the big curve; one loaded truck about to enter the curve before another loaded truck

Table 2 Performance for alternative configurations and strategies.

Number of Trucks Used in Operation	*Minimum Clear Distance "d" to Prevent Loaded Trucks From Entering Curve*	*Minimum Number of Empty Trucks "n" Waiting to Enter Curve to Prevent Entry of Loaded Trucks*	*Strategy to Route Trucks to Excavators*	*Production (lcm/hr)*	*Main Excavator Utilization*	*Alternate Excavator Utilization*	*Average Time Trucks Wait for Main Excavator (min)*	*Average Time Trucks Wait for Alternate Excavator (min)*	*Average Time Loaded Trucks Wait to Enter Curve (min)*	*Average Time Empty Trucks Wait to Enter Curve (min)*
12	470	1	MC(66.7%)	764.3	61.8%	46.4%	3.19	2.23	0.06	2.15
12	470	1	Dyn(1.566,2.066)	794.3	62.8%	50.3%	3.18	1.70	0.03	1.65
12	0	1	Dyn(1.566,2.066)	823.8	87.5%	17.8%	1.46	0.13	1.76	0.80
12	0	2	Dyn(1.566,2.066)	887.6	89.3%	26.6%	1.32	0.06	0.70	1.20
12	0	3	Dyn(1.566,2.066)	868.1	84.7%	30.0%	1.50	0.08	0.50	1.61
12	260	1	Dyn(1.566,2.066)	901.2	96.3%	17.7%	1.46	0.10	0.74	0.85
13	263	1	Dyn(1.566,2.066)	950.3	96.5%	26.7%	1.60	0.11	0.76	0.93
14	251	1	Dyn(1.566,2.066)	999.7	97.5%	34.6%	1.73	0.11	0.81	0.97
15	250	1	Dyn(1.566,2.066)	1046.9	98.0%	42.6%	1.91	0.12	0.84	1.01
16	182	1	Dyn(1.566,2.066)	1091.2	99.3%	48.9%	2.16	0.12	0.93	0.95
16	152	1	TAM(9)	1110.1	80.7%	82.0%	0.88	0.81	0.93	0.95
16	145	1	TAM(10)	1114.8	88.1%	71.3%	1.07	0.56	0.99	0.94
16	161	1	TAM(11)	1112.4	94.8%	60.3%	1.41	0.35	0.99	0.94
16	212	1	TAM(12)	1092.9	99.2%	49.6%	2.04	0.18	0.95	0.97
16	142	1	Dyn(1.15,0.9)	1120.7	94.7%	61.9%	1.31	0.24	1.01	0.94

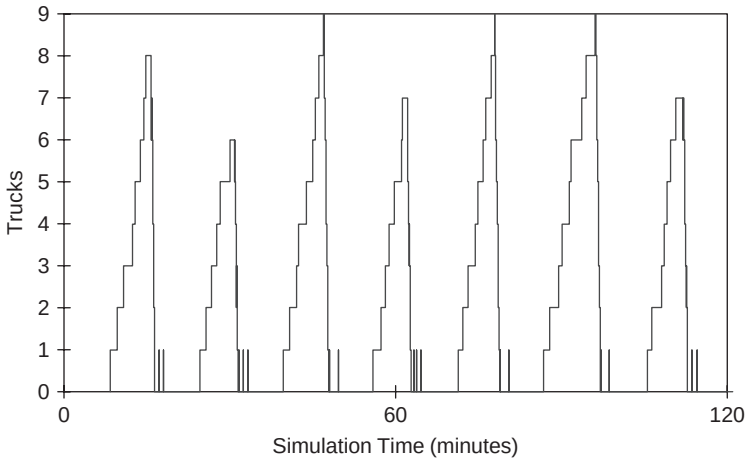


Figure 3 Trucks waiting to enter curve empty in initial design.

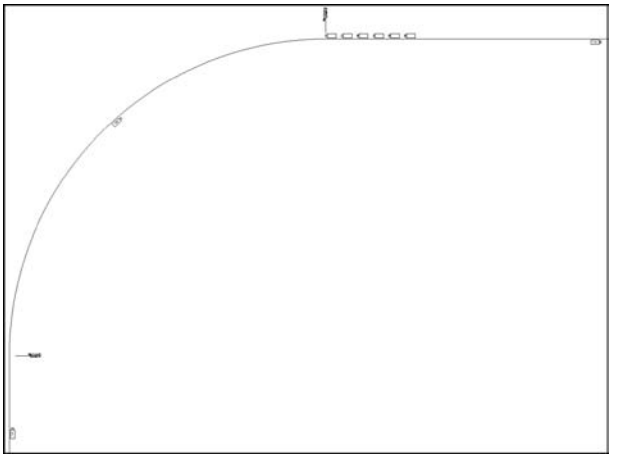


Figure 4 Animation snapshot showing bunched up trucks waiting to enter the narrow curve.

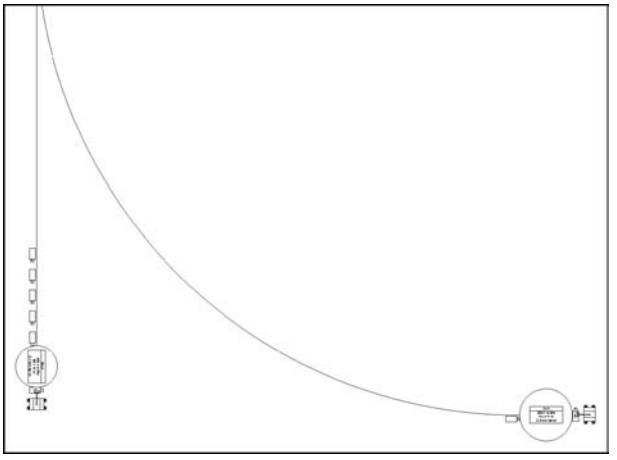


Figure 5 Animation snapshot showing bunched up trucks waiting for excavators.

finishes traversing it; and one loaded truck heading towards the dump area. Two trucks are out of sight at the dump area and will soon join the bunch waiting to enter the curve.

The slow speed of the loaded trucks and the fast speed of the empty trucks as they traverse the curve cannot be seen on the snapshot. Only the animation can convey that information. Due to the slow speed and long separation, loaded trucks often enter the curve when another loaded truck has almost passed it completely. This holds the empty trucks for long periods of time during which they bunch up.

Figure 5 shows a snapshot of the animation five minutes later, when the trucks that had bunched up have arrived almost together to the loading areas. Three trucks are out of view in the snapshot but will arrive soon to the loading areas.

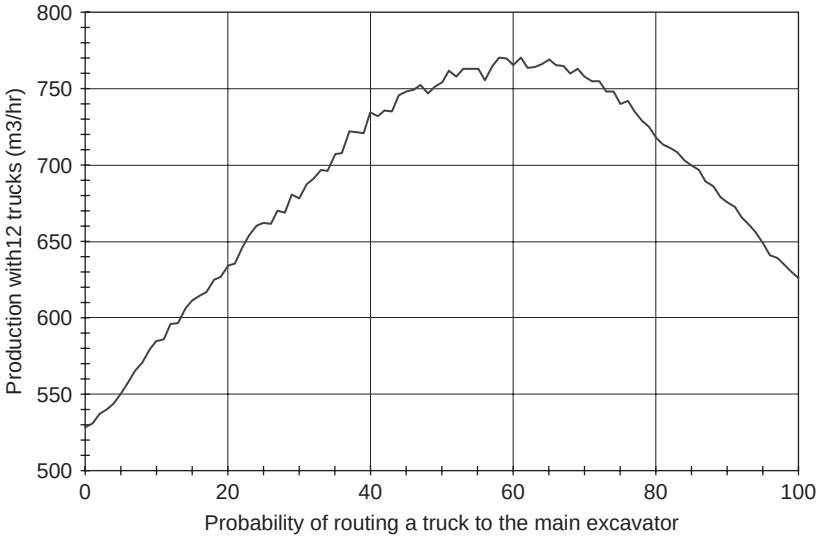


Figure 6 Production as a function of the percentage of trucks routed to the main excavator using a total of 12 trucks.

The animation also reveals the ineffectiveness of the probabilistic (but random) truck routing policy; at times, trucks are routed to an excavator that is busy and has a long queue of trucks waiting to be served even though the other excavator is free. The 66.7% probability used to route to the main area and the 33.3% probability used to route to the alternate area were determined rather naively, under the assumption that the main excavator should be used as much as possible. Figure 6 shows the production rates that could be achieved using different routing percentages. This indicates that the percentages used in the preliminary design were not a bad choice, although the routing method itself is.

6 Improvements to the operation

The operation needs improvements in two major areas: the strategies for routing trucks to the loading areas and for establishing direction of travel in the narrow curve. Both aspects are interrelated. The optimum method for setting travel direction in the curve combined with a given truck routing strategy may not be optimal with a different routing strategy. The converse may also happen. As a first step it appears convenient to design a simple but scaleable truck routing strategy that will work regardless of the number of trucks used in the operation or on how the narrow curve is administered.

One such concept is that of minimum turnaround time to the point of decision. In this case, the point of decision is the vertex that joins *LdRts*, *MLR*, and *ALR* (for trucks traveling on *LdRts* towards the loading areas). For each loading area, this strategy requires a forecast of the time required for a truck to get there, load, and return. The quickest option (i.e., the option that allows the truck to be back with a load faster) is then selected. In forecasting the turnaround time for a truck, it is necessary to estimate

the expected time to the beginning of the actual truck loading. This is the largest of 1) the time required to travel to the excavator, and 2) the time for the excavator to complete the current loading plus the loading of any other trucks ahead of the one being considered. The output of the initial design indicated an average time to travel to the main shovel of 0.52 minutes, average time to position under the excavator of 0.17 minutes and average loading time of 1.05 minutes; and an average travel time to the alternate shovel of 0.93 minutes, average time to position under the excavator of 0.18 minutes and average loading time of 1.73 minutes. Assuming that excavators are half way through a loading at the time of the routing decision, then the time to travel to the excavator is very close to the remaining loading time and can be disregarded. Thus, a simple estimate of the time to the beginning of service is (number of trucks ahead of the one being routed and waiting to load) * (average time to position and load each truck) + (half the time of the assumed current loading). To obtain the complete turnaround time it is necessary to add the time to position and load the truck being routed and the time to exit the load area and return to the point of decision (1.35 minutes to exit the load area and return from the main and 3.53 minutes from the alternate). These minimum turnaround times can be expressed as follows:
To Main:

$$1.22^{*}(\text{MLRnTrucksTrvl} + \text{MLAnTrucksWt}) + 0.53 + 1.35 \tag{2}$$

To Alternate :

$$1.91^{*}(\text{ALRnTrucksTrvl} + \text{ALAnTrucksWt}) + 0.87 + 3.53 \tag{3}$$

In the above expressions, *MLRnTrucksTrvl* and *ALRnTrucksTrvl* are the number of trucks currently travelling on *MLR* and *ALR* (respectively) towards the excavators. *MLAnTrucksWt* and *ALAnTrucksWt* are the number of trucks waiting to load at the main and alternate loading areas (respectively). These expressions are dynamic because they access information about the constantly changing state of the system. These expressions can be combined and simplified into a logical expression that is true only when the turnaround time to the main excavator is smaller than that to the alternate excavator as follows:
Main Preferred:

$$(\text{MLRnTrucksTrvl}+\text{MLAnTrucksWt}) <= 1.566^{*} (\text{ALRnTrucksTrvl} + \text{ALAnTrucksWt}) + 2.066 \tag{4}$$

By reversing the comparison it is also possible to indicate that the turnaround time to the alternate excavator is the smallest: Alternate Preferred:

$$(\text{MLRnTrucksTrvl} + \text{MLAnTrucksWt}) > 1.566^{*} (\text{ALRnTrucksTrvl} + \text{ALAnTrucksWt}) + 2.066 \tag{5}$$

When these expressions are used to route trucks nothing is left to chance – everything is calculated based on a very specific rule.

The results obtained by implementing this change in the truck routing strategy are shown in the second row of Table 2 (production of 794.3 lcm/hr). This new rule increases production by 3.9%. This improvement is small because the narrow curve remains the main bottleneck. Trucks still bunch up when attempting to enter the curve empty.

6.1 *Managing the curve*

The basic problem with the narrow curve is that loaded trucks, traveling uphill, are very slow. They arrive at the curve at such an interval that they enter when a previous truck is almost exiting. The direction of travel is thus maintained in the loaded direction for very long periods, during which, empty trucks arrive and bunch at the other end. When empty trucks enter the curve, however, they traverse and clear it very quickly. The entry of loaded trucks to the curve must be controlled so that empty trucks can traverse it in smaller bunches.

An obvious and simple strategy might be to prevent loaded trucks from entering the curve when n empty trucks are waiting to enter at the other side. When the model is modified in this way, the results obtained with $n = 1$, 2 and 3 are shown in rows 3, 4, and 5 of Table 2, with productions of 823.8, 887.6, and 868.1 lcm/hr. With $n = 2$, production increases to 82% of project requirements. The waiting times are more evenly distributed, and the trucks bunch up in groups of 2, 3, and 4 instead of groups of up to 12. **This shows that even planning on the basis of these preliminary improvements would have exposed the contractor to significant risks in being able to complete the entire project on time and within budget, had a simulation analysis not been performed.**

Although a good improvement, further improvements may be possible. Sometimes a loaded truck arrives to the curve with the previous truck just a short distance into the curve. The truck is prevented from entering because n trucks are already waiting at the other end. In these cases, it seems that allowing the truck to enter may be a good idea, as it will only delay the entrance of empty trucks by a short time. In order to consider this, the strategy for managing the narrow curve may be modified to prevent entry of loaded trucks when n empty trucks are waiting at the other end **and** the previous truck is into the curve by a distance larger than d. With $n = 1$ and $d = 100$ meters, for example, loaded trucks are prevented from entering the curve if at least 1 empty truck is waiting at the other end and if the previous truck is further than 100 meters into the curve.

Figure 7 shows the resulting production for $n = 1$, 2 and 3 as a function of d. The effect of d for $n = 2$ and 3 is hardly noticeable for values of d smaller than 200 and 260 meters, respectively. The impact of d for $n = 1$ corresponds with the motivation for implementing this strategy. The results for $n = 1$ and $d = 260$ meters are shown in row 6 of Table 2. The production is 901.2 lcm/hr, about 1.5% higher than the production obtained with $n = 2$ and $d = 200$ meters. In every case tested, $n = 1$ outperformed $n = 2$ when each was adjusted to the most convenient d. Notice that $d = 0$ meters is a special case that is equivalent to not considering d at all. Similarly, $d = 470$ meters is a special case equivalent to not considering n or d at all. **Even with these modifications, the contractor would have faced difficulty in completing the project on time and in budget, had it not been for the simulation analysis.**

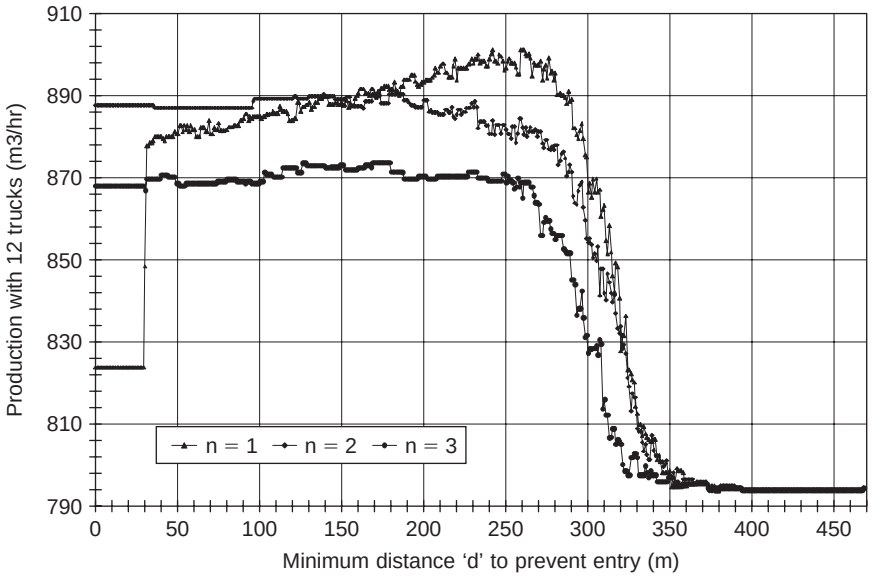


Figure 7 Production for n = 1, 2, and 3 as function of *d* using 12 trucks.

6.2 *Reaching the required production*

Although further improvements are possible, it is very unlikely that the required 1083 lcm/hr will be reached without increasing the number of trucks.

Figure 8 shows the production as a function of d for $n = 1$ and 12 to 16 trucks. The results at the respective optimum values of d are summarized in rows 6 to 10 of Table 2. Sixteen trucks with $n = 1$ and $d = 182$ meters are required to provide 1091.2 lcm/hr and surpass the minimum required production of 1083 lcm/hr.

Other strategies could be explored in setting the direction of travel for the curve, perhaps based on the number of trucks on each side or on a fixed direction pattern such as the one used by standard traffic lights. The improvements achieved here, however, seem appropriate to illustrate the idea that the curve's logistics have a significant impact in production and must be designed with care.

6.3 *Fine tuning the truck routing strategy*

The truck routing strategy, however, can be refined with the hopes of increasing production and perhaps lowering the load on the main excavator.

6.3.1 *Assigning specific trucks to specific shovels*

One routing method in use by some contractors is to assign a specific excavator to each truck. Each truck goes to its excavator even if several other trucks are waiting for that excavator and other excavators are free. An investigation of the impact on production of this simple and practical approach yielded the surprising results shown in Figure 9.

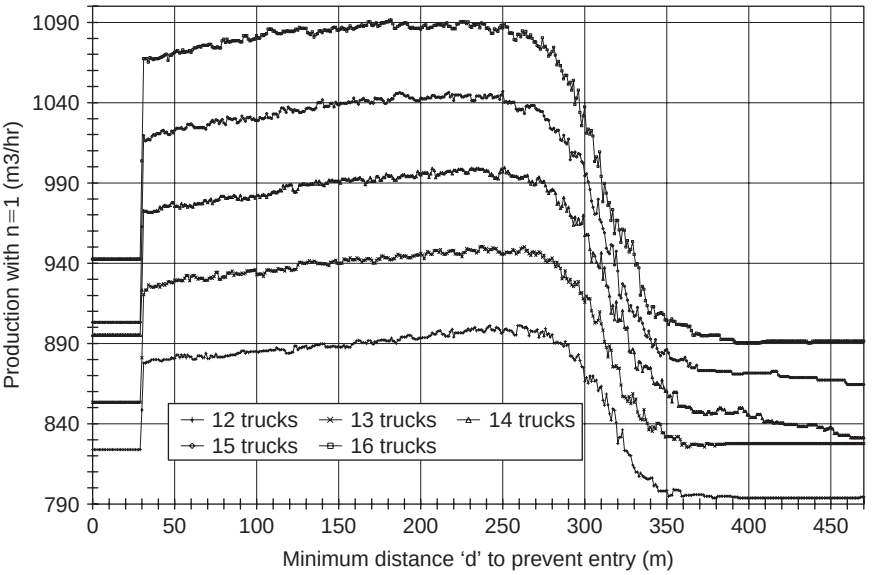


Figure 8 Maximum production and corresponding values of n and d with simple dynamic truck routing.

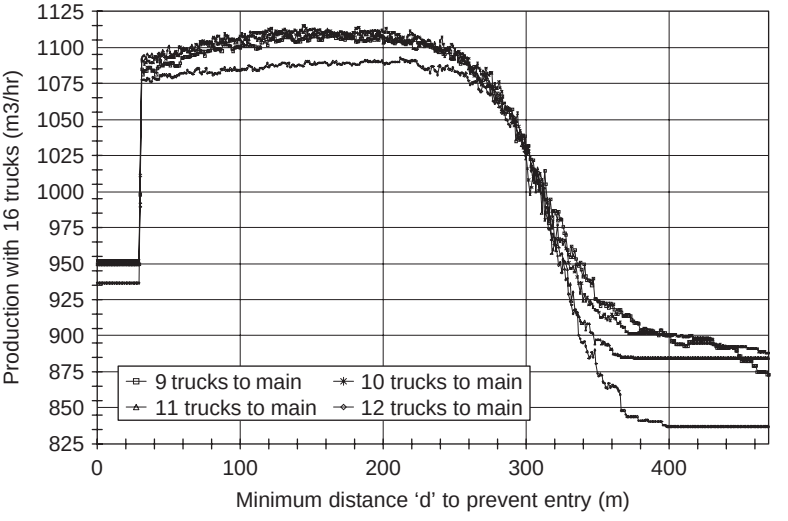


Figure 9 Production as function of *d* with a specific number of the 16 trucks assigned to the main excavator.

In this figure, the production rates as a function of *d* are plotted for the cases where 9, 10, 11, and 12 trucks are assigned to the main excavator. The maximum points of each curve are reported in rows 11 to 14 of Table 2, with production rates of 1110.1, 1114.8, 1112.4, and 1092.9 lcm/hr, respectively. These production rates exceed the one

achieved with the previous dynamic strategy. The reasons for the increased production rates achieved by assigning trucks to shovels for the entire operation are difficult to pinpoint, and probably have to do with the regularity imposed.

6.3.2 *Better dynamic truck routing*

Countless truck routing strategies can be explored. These strategies may vary in their sophistication and in the amount of state-dependent data that they consider. The dynamic strategy explored earlier (minimum turnaround) was based on linear equations that considered the number of trucks traveling towards, and waiting to load at, each loading area. Other dynamic data was assumed. Data of this type include 1) the time already spent by the shovel in the current loading if a loading is currently taking place (the prior assumption was that a loading was indeed taking place and that it was half done); and 2) if there are no trucks waiting to load, but at least one is on the way, the time at which the most advanced of them is expected to arrive at the loading area. Including such dynamic data in a model is possible, but dynamically determining this data while the operation is taking place in the field is not practical.

It is possible, however, to capture some of the impact of that data by using fixed deterministic values for them that are different from those assumed. Retaining a linear equation for the decision, any changes in assumed values will show up in the slope and intercept of the linear equation. Thus, the previous dynamic routing criteria determined by expression (4) can be generalized as

$$(\text{MLRnTrucksTrvl} + \text{MLAnTrucksWt}) <= \textit{Multiplier}^{*} (\text{ALRnTrucksTrvl}+\text{ALAnTrucksWt}) + \textit{Adder} \quad (6)$$

Expression (5) can also be generalized by replacing the '<=' in expression (6) with '>'. By trying different values of *Multiplier* and *Adder*, it may be possible to capture some aspect of the routing that cannot be readily conceptualized analytically, but that may yield better performance. Setting *Multiplier* to 1.15 and *Adder* to 0.9 (values obtained by brute force), for example, yields a production rate of 1120.7 lcm/hr with 16 trucks, $n = 1$, and $d = 142$ meters (row 15 in Table 2). This production exceeds those obtained by assigning trucks to specific excavators. Figure 10 shows production with 16 trucks as function of d, for the three different routing strategies: 1) the original dynamic case with *Multiplier* set to 1.566 and *Adder* to 2.066, 2) assigning 10 trucks to the main excavator, and 3) the dynamic case with *Multiplier* set to 1.15 and *Adder* set to 0.9.

Even with 16 trucks and sophisticated strategies for administering the narrow curve and routing trucks, the constraints imposed by the site limit production to 1120.7 lcm/hr. This is lower than the 1166 lcm/hr that were estimated initially based on only 12 trucks.

6.3.3 *Impact of uncertainty*

Variability impacts most construction field operations. The example here is not different. The deterministic run of the alternative in row 15 of Table 2, for example, yields a production of 1142.2 lcm/hr. This production is higher than the 1120.7 lcm/hr obtained when uncertainty is considered. The difference, however, is very small. In

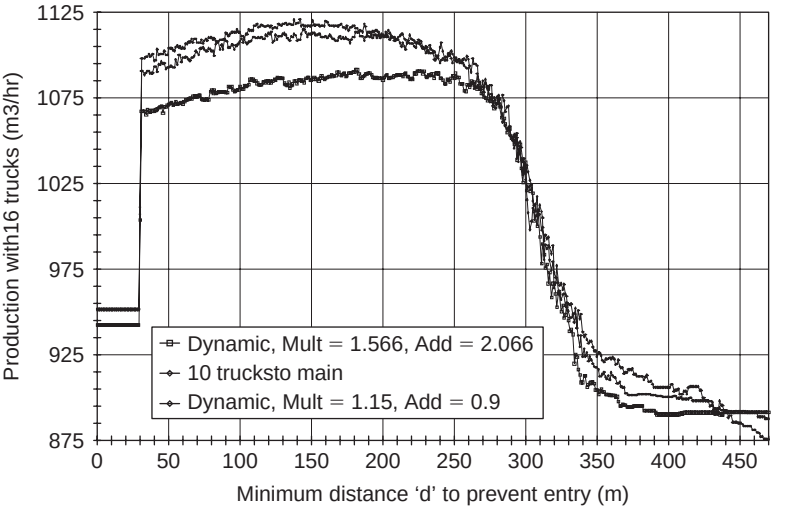


Figure 10 Production for various truck routing strategies as function of *d* using 16 trucks and n = 1.

cases like the one presented in this paper, the impact of operating strategy is much more significant than the impact of uncertainty. This indicates that proper modeling of strategy is far more important than obtaining accurate probability distributions for the different aspects of this operation.

6.3.4 *Field implementation of routing strategy*

In order to implement the best of the strategies discussed, it is necessary to use sensors and electronic signals. Sensors could be placed on the haul routes to keep track of how many trucks are in the different segments or areas. Large electronic signals could be linked to the sensors, make the appropriate calculations, and indicate to truck operators the load area to which they should go. Such a mechanism may be justified only for large production volumes and has been implemented in several mines (Chironis 1985; Lizotte and Bonates 1987).

In practice, however, it may be more convenient to assign trucks to specific excavators. This does not require sophisticated technology and can be implemented very simply. The production levels attained are not the highest possible, but the convenience of simplicity far outweighs the minimal decrease in performance compared to the optimal dynamic case.

7 Conclusion

The design of a complex earthwork operation involves, in addition to the selection of methods, equipment, and haul routes, the adoption of dynamic strategies that often play a significant role in performance. Discrete-event simulation is a method of analysis that is capable of considering the complexities and strategies of such operations.

Implementing dynamic strategies in the virtual world modeled by a discrete-event tool is easier than implementing them in the field. The dynamic truck routing strategies described in this paper are a clear example. The truck routing also illustrates that some common practices, although simple, may perform well when compared with complex sophisticated strategies. Assigning trucks to specific loading areas is a clear example that works that way, at least in the case study presented in this paper. Traditional methods do not always perform that well. Each case should be evaluated based on the specific site conditions and fleet configuration details that apply to it.

Dynamic strategies are not exclusive to earthwork operations; they are present in many types of construction processes. Modeling such strategies brings a level of realism to simulations that may contribute to an acceptance of the methodology by practitioners, who otherwise may not use it due to a lack of awareness of the fact that their complex ideas can indeed be modeled. Recent advances in discrete-event simulation place the modeling of such strategies within reach.

By using discrete-event simulation to analyze and design construction operations, it is possible for the contractor to reduce his exposure to the risks associated with not meeting project milestones and enduring excessive unanticipated costs.

Acknowledgements

The author thanks the National Science Foundation (Grant CMS-9733267) for supporting portions of the work presented here. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Chironis, N.P. (1985). "Computer Monitors and Controls all Truck-Shovel Operations", Coal Age, 90(03), pp 50–55.

Gransberg, D.D. (1996). "Optimizing Haul Unit Size and Number Based on Loading Facility Characteristics," Journal of Construction Engineering and Management, ASCE, 123(3), pp 248–253.

Ioannou, P.G., and Martinez, J.C. (1996). "Comparison of Construction Alternatives Using Matched Simulation Experiments," Journal of Construction Engineering and Management, ASCE, (112) 3, New York, NY.

Law, A.M. and W.D. Kelton (2000). Simulation Modeling and Analysis, 3rd Edition. McGraw Hill, New York, NY, USA.

Lizotte, Y., and Bonates, E. (1987). "Truck and Shovel Dispatching Rules Assessment Using Simulation," Mining Science and Technology, 5, pp 45–58.

Martinez, J.C. (1996). "STROBOSCOPE: State and Resource Based Simulation of Construction Processes." PhD Dissertation, University of Michigan, Ann Arbor, MI.

Martinez, J.C. (1998). "Earthmover-Simulation Tool for Earthwork Planning", Proceedings of the 1998 Winter Simulation Conference, pg. 1263–1271, vol 2.

Schexnayder, C., Weber, S.L., and Brooks, B.T. (1999), "Effect of Truck Payload Weight on Production," Journal of Construction Engineering and Management, ASCE, 125(1), pp 1–7.

References

AASHTO. 2004. *LRFD bridge design specifications*. Washington, D.C.: American Association of State Highway and Transportation Officials (AASHTO).

Adeli, H. and Kamal, O. 1986. Efficient Optimization of Space Trusses, *Computers and Structures*, 24(3), 501–511.

Ahmad, S. 2003. Reinforcement Corrosion in Concrete Structures, its Monitoring and Service Life Prediction – A Review, *J. of Cement and Concrete Composites*, No. 25, Elsevier Press, pp. 459–471.

AISC. 1989. *Manual of Steel Construction – Allowable Stress Design*, American Institute of Steel Construction, Chicago, Illinois.

Albrecht, P. 1983. S-N Fatigue Reliability Analysis of Highway Bridges, Probabilistic Fracture Mechanics and Fatigue Methods: Application for Structural Design and Maintenance, ASTM STP 798.

Albrecht, P.A. and Naeemi, A.H. 1984. Performance of weathering steel in bridges. *NCHRP Report 272*. Washington, D.C.: Transportation Research Board.

Albrecht, P.A., Coburn, F.M., Wattar, G.L., Tinklenberg, G.L. and Gallagher, W.P. 1989. Guidelines for the use of weathering steel in bridges. *NCHRP Report 314*. Washington, D.C.: Transportation Research Board.

Aldunate, R., Ochoa, S.F., Pena-Mora, F., and Nussbaum, M. 2006. Robust Mobile Ad Hoc Space for Collaboration to Support Disaster Relief Efforts Involving Critical Physical Infrastructure, *Journal of Computing in Civil Engineering*, Vol. 20, pp. 13–27.

Allen, D.E. 1990. Building Vibrations from Human Activities. *Concrete International Design and Construction* 12:6, 66–73.

Allen, T.M. and Holtz, R.D. 1991. Design of Retaining Walls Reinforced with Geosynthetics, State of the Art Paper, session 8A on Earth Reinforcement, *Geotechnical Engineering Congress 1991*, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 970–987.

Allen, T.M., Bathurst, R.J., Holtz, R.D., Walters, D. and Lee, W.F. 2003. A New Working Stress Method for Prediction of Reinforcement Loads in Geosynthetic Walls, *Canadian Geotechnical Journal*, Vol. 40, No. 5, pp. 976–994.

Allen, T.M., Bathurst, R.J., Lee, W.F., Holtz, R.D. and Walters, D. 2004. A New Working Stress Method for Prediction of Reinforcement Loads in Steel Reinforced MSE Walls, *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 130, No. 11, pp. 1109–1120.

American Association of State Highway and Transportation Officials (AASHTO). 1998. *LRFD Bridge Design Specifications*, 2nd edition, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). 2007. *LRFD Bridge Design Specifications*, 2nd edition, Washington, D.C.

Amick, H. and Bayat, A. 1998. Dynamics of Stiff Floors for Avanced Technology Facilities. *Proceedings of 12th ASCE Engineering Mechanics Conference*, La Jolla, CA, 318–321,

Amick, H., Hardash, S., Gillett, P. and Reaveley, R.J. 1991. Design of Stiff, Low-Vibration Floor Structures. *Proceedings of International Society for Optical Engineering (SPIE)*, San Jose, CA, Vol. 1619, 180–191.

Ang, A.H-S. 2007. An application of quantitative risk assessment in infrastructures engineering. *Proceedings CECAR 4*, Taipei, Taiwan.

Ang, A.A-S. and Tang, W.H. 1984. Probability Concepts in Engineering Planning and Design: Volume II – Decision, Risk and Reliability, J. Wiley & Sons, Inc., New York.

Ang, A.H-S. and Tang, W.H. 1990. Probability Concepts in Engineering Planning and Design. Volume II- Decision, Risk, and Reliability. John Wiley & Sons, Inc. New York.

Ang, A.A-S. and Tang, W.H. 2007. Probability Concepts in Engineering, 2nd Edition, J. Wiley & Sons, Inc., New York.

Ang, A.H.-S. and Tang, W.H. 2007. *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, Vol. 1, 2nd edition, Wiley, New York.

Applied Technology Council. 1985. *Tentative Provisions for the Development of Seismic Regulations for Buildings, ATC-3-06*, National Bureau of Standards, Washington D.C., 1978; also BSSC.

AREMA. 2005. *Manual of railway engineering*. Washington, D.C.: American Railway Engineering and Maintenance of Way Association (AREMA).

Arora, J.S. 2004. *Introduction to Optimum Design*, 2nd Edition, El-Sevier Inc., California.

ASCE, 2005. Report card for America's Infrastructure, American Society of Civil Engineers, Reston, VA. Available online at www.asce.org/reportcard/2005/index.cfm (accessed 22 May 2007).

ASCE Standard 43-05. 2005b. *Seismic design criteria for structures, systems and components in nuclear facilities*. American Society of Civil Engineers, Reston, VA.

ASCE. Structural Health Monitoring Committee. 2000. Benchmark studies. Available online at http://cive.seas. wustl.edu/wusceel/asce.shm/benchmarks.htm (accessed June 2007).

ASCE/ANSI Standard 7-05. 2005a. *Minimum design loads for buildings and other structures*. American Society of Civil Engineers, Reston, VA.

ASCE/ANSI Standard 41-06. 2006. *Guidelines for the Seismic Rehabilitation of Buildings* (previously published as FEMA 356), American Society of Civil Engineers, Reston, VA.

ASTM C876-91. 1991. Standard Test Method for Half-Cell Potential of Reinforcing Steel in Concrete, *American Society for Testing Materials*, 03(02), pp. 434–439.

Ayyub, B.M. 2001. Elicitation of Expert Opinions for Uncertainty and Risks, CRC Press, Boca Raton, FL.

Ayyub, B.M. 2003. Risk Analysis in Engineering and Economics, Chapman and Hall/CRC Press, FL.

Ayyub, B.M. and McCuen, R. 2003. Probability, Statistics and Reliability for Engineers and Scientists, Second Edition, Chapman and Hall/CRC Press, Boca Raton, FL.

Ayyub, B.M. and Klir, G.J. 2006. Uncertainty Modeling and Analysis for Engineers and Scientists, Chapman & Hall/CRC, Press Boca Raton, FL.

Ayyub, B.M. and McGill, W.L. 2008. An All-Hazards Methodology for Ciritical Asset and Portfolio Risk Analysis, ASCE Special Publication *Primer on Quantitative Risk Analysis* (in press).

Ayyub, B.M. Foster, J. and McGill, W.L. 2007a. Risk Analysis of a Protected Hurricane-Prone Region: Model Development, ASCE Natural Hazards Review, in press.

Ayyub, B.M., McGill, W.L., Foster, J. and Jones, H.W. 2007b. Risk Analysis of a Protected Hurricane-Prone Region: Computations and Illustrations, ASCE Natural Hazards Review, in press.

Ayyub, B.M., McGill, W.L. and Kaminskiy, M. 2007. Critical Asset and Portfolio Risk Analysis for Homeland Security: An All-Hazards Framework, Risk Analysis International Journal, Society for Risk Analysis, 27(3), 2007, 789–801.

Bachmann, H. and Ammann, W. 1987. Vibrations in Structures Induced by Man and Machine. IABSE-AIPC-IVBH, Vol. 3e. International Association for Bridge and Structural Engineering (IABSE), Zurich, Switzerland, 1–43.

Balaguru, P. 1998. Geopolymer for Protective Coating of Transportation Infrastructures, *Technical Report, CAIT/Rutgers, the* State *University*, New Jersey.

Bathurst, R.J. and Jones, C.J.F.P. 2001. Chapter 17 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 501–537.

Bathurst, R.J., Walters, D., Vlachopoulos, N., Burgess, P. and Allen, T.M. 2000. Full Scale Testing of Geosynthetic Reinforced Walls, *Advances in Transportation and Geoenvironinental Systems Using Geosynthetics*, Proceedings of sessions at GeoDenver 2000, J.G. Zornberg and B.R. Christopher, Editors, Geotechnical Special Publication No. 103, ASCE, pp. 201–217.

Bazant, Z.P. 1979. Physical Model for Steel Corrosion in Concrete Sea Structures – Application, *ASCE J. Struct. Div.*, 105(6), 1155–1166.

Bell, K.R., Davie, J.R., Clemente, J.L. and Likins, G. 2002. Proven Success for Driven Pile Foundations, *Proc. Int. Deep Foundations Congress, Geotech. Spec. Publ. No. 116* (ed. by O'Neill, M.W. and Townsend, F.C.), ASCE, Reston, Va., pp. 1029–1037.

Benjamin, J.R. and Cornell, C.A. 1970. Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill, Inc., New York.

Berg, R.R. 1993. *Guidelines for Design, Speci fication, & Contracting of Geosynthetic Mechanically Stabilized Earth Slopes on Firm Foundations*, Federal Highway Administration, Report No. FHWA-SA-93-025, 87 pp.

Berkeley, K.G.C. and Pathmanaban, S. 1990. *Cathodic Protection of Reinforcement Steel in Concrete*, Butterworth-Heinemann, London.

Bertolini, L., Elsener, B., Pedeferri, P. and Polder, R. 2004. *Corrosion of Steel in Concrete*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.

Bijen, J. 2003. *Durability of Engineering Structures: Design, Repair and Maintenance*. Woodhead Publishing Limited, Abington Hall, Cambridge, England.

Bonaparte, R. and Christopher, B.R. 1987. Design and Construction of Reinforced Embankments over Weak Foundations, *Transportation Research Record 1153*, TRB, pp. 226–239.

Bonaparte, R., Holtz, R.D. and Giroud, J.P. 1987. Soil Reinforcement Design Using Geotextiles and Geogrids, in *Geotextile Testing and the Design Engineer, ASTM Special Technical Publication 952*, J.E. Fluet, editor, American Society for Testing and Materials, pp. 69–116.

Boore, D.M. and Joyner, W.B. 1994. Prediction of ground motion in North America. ATC-35-1, 6-1 to 6-41.

Boore, D.M., et al. 2003. Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard Bridge collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America*, **93**:2737–2751.

Brad, A.J. and Faulkner, L.R. 2001. *Electrochemical Methods Fundamentals and Applications*, 2nd edition, John Wiley & Sons, New York.

Brockwell, P.J. and Davis, R.A. 2002. *Introduction to Time Series and Forecasting*. Springer-Verlag, Second Edition, New York.

Broomfield, J.P. 1997. *Corrosion of Steel in Concrete*, E&FN Spon, London.

Bruneau, M., Chang, S.E., Eguchi, R.T., Lee, G.C., O'Rourke, T.D., Reinhorn, A.M., Shinozuka, M., Tierney, K., Wallace, W.A. and von Winterfeldt, D. 2003. A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra*, Vol. 19, No. 4, pp. 733–752.

Bucher, C. and Frangopol, D.M. 2006. Optimization of lifetime maintenance strategies for deteriorating structures considering probabilities of violating safety, condition, and cost thresholds. *Probabilistic Engineering Mechanics*, Elsevier, 21(1), 1–8.

Budelmann, H. and Hariri, K. 2006. A Structural Monitoring System for RC-/PC-Structures, *Proceedings LCC5*, Seoul, Korea.

Canadian Highway Bridge Design Code (CHBDC). 2006. Ministry of Transportation, Ontario, Canada.

Casas, J.R., Frangopol, D.M. and Nowak, A.S., eds., 2002. *Bridge Maintenance, Safety and Management*, CIMNE, Barcelona, Spain.

Catbas, F.N., Susoy, M. and Frangopol, D.M. 2008. Structural Health Monitoring and Reliability Estimation: Long Span Truss Bridge Application with Environmental Monitoring Data, *Engineering Structures*, Elsevier (in press).

Catbas, N.F., Zaurin, R., Frangopol, D.M. and Susoy, M. 2007. System Reliability-Based Structural Health Monitoring with FEM Simulation. *Proceedings SHMII-3*, Vancouver, B.C., Canada.

CEN. 2001. *Eurocode 7-Geotechnical Design, Part 1 General Rules*, European Committee for Standardization (CEN).

Chan C.M. 1997. How to Optimize Tall Steel Building Frameworks, *ASCE Manuals and Report on Engineering Practice No. 90 – Guide to Structural Optimization*, ASCE, 165–196.

Chang, C.C., Ger, J.R. and Cheng, F.Y. 1994. Reliability-Based Optimum Design for UBC and Nondeterministic Seismic Spectra, *Journal of Structural Engineering*, ASCE, Vol. 120, No. 1, pp. 139–160.

Chang, F-K. (ed.) 1999, 2001, 2003, 2005, 2007 and, *1st, 2nd, 3rd, 4th and 5th International Workshops on Structural Health Monitoring*, Stanford University, Stanford, CA. 94305.

Chang, S.E. and Seligson, H.A. 2003. Evaluating Mitigation of Urban Infrastructure Systems: Application to the Los Angeles Department of Water and Power. in *Proceedings of the 2003 Technical Council on Lifeline Earthquake Engineering Conference*, Reston, Virginia.

Chang, S.E., Shinozuka, M. and Moore, J. 2000. Probabilistic Earthquake Scenarios: Extending Risk Analysis Methodologies to Spatially Distributed Systems. *Earthquake Spectra*, Vol. 16, No. 3, August, pp. 557–572.

Chang, S.E., Svekla, W.D. and Shinozuka, M. 2002. Linking Infrastructure and Urban Economy: Simulation of Water Disruption Impacts in Earthquake. *Environment and Planning B*, Vol. 29, No. 2, pp. 281–301.

Cheng, F.Y. (ed.). 1986. *Recent Developments in Structural Optimization*, ASCE.

Cheng, F.Y. 1986. *Optimum Design of Reinforced Concrete and Steel Structures Subjected to Static and Seismic Loads*, National Taiwan Institute of Technology, Taiwan, Vols. I (242 pages) and II (375 pages).

Cheng, F.Y. 2001. *Matrix Analysis of Structural Dynamics – Applications and Earthquake Engineering*, Marcel Dekker, Inc., also CRC Press, Taylor and Francis Group, Boca Raton, Florida.

Cheng, F.Y. and Botkin, M.E. 1976. Nonlinear Optimum Design of Dynamic Damped Frames, *J. of Structural Div.*, ASCE., Vol. 102, No. ST3, pp. 609–628, March.

Cheng, F.Y. and Chang, C.C. 1988. *Safety-Based Optimum Design of Nondeterministic Structures Subjected to Various Types of Seismic Loads*, NSF Report, U.S. Department of Commerce, National Technical Information Service, Virginia, NTIS No. PB90-133489/AS (326 pages).

Cheng, F.Y. and Juang, D.S. 1988. Assessment of Various Code Provisions Based on Optimum Design of Steel Structures, *Journal of Earthquake Engineering and Structural Dynamics*, Vol. 16, pp. 46–61.

Cheng, F.Y. and Li, D. 1997. Multiobjective Optimization Design with Pareto Genetic Algorithm, *Journal of Structural Engineering*, ASCE, Vol. 123, No.9, pp. 1252–1261.

Cheng, F.Y. and Srifuengfung, D. 1978. Optimum Structural Design for Simultaneous Multicomponent Static and Dynamic Input, *International Journal for Numerical Methods in Engineering*, Vol. 13, No.2, pp. 353–372.

Cheng, F.Y., Ang, A.H-S., Lee, J.R. and Li, D. 1999. Genetic Algorithm for Multiobjective Optimization and Life-Cycle Cost, *Structural Engineering in the 21st Century*, ASCE, pp. 484–489.

Cheng, F.Y., Hong, H.P. and Lou, K.Y. 2008. *SMART STRUCTURES – Innovative Systems for Seismic Response Control*, CRC Press, Taylor and Francis Group, Boca Raton, Florida.

Cheng, F.Y., Venkayya, V.B. and Khachaturian, N., et al. 1976. *Computer Methods of Optimum Structural Design*, Vols. 1 and 2, University of Missouri-Rolla, Rolla, Missouri.

Chironis, N.P. 1985. Computer Monitors and Controls all Truck-Shovel Operations, Coal Age, 90(03), pp 50–55.

Cho, H.N. 2008. Practical Approaches to the Application of Life-Cycle Civil Engineering for Infrastructures in Korea, *International Association Life-Cycle Civil Engineering'08, Keynote*.

Cho, H.N., Ang, A.H-S. and Lim, J.K. 1999. Cost-effective optimal target reliability for seismic design and upgrading of long span PC bridges. *Proceedings of the 8th International Conference on Applications of Statistics and Probability.,* Newcastle, Australia: Department of Civil, Surveying and Environmental Engineering, University of Newcastle.

Cho, H.N., Choi, H.H. and Kim, Y.B. 2002. A risk assessment methodology for incorporating uncertainties using fuzzy concepts. *Reliability Engineering and System Safety*. 78(2): 173–183.

Cho, H.N., Choi, H.H., Lee, K.M. and Park, K.H. 2006. Optimal seismic retrofit and maintenance strategy for steel bridges using Life-Cycle Cost analysis. *IABMAS. Proceedings of the 5th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems,* Seoul, Korea.

Cho, H.N., Kim, J.H., Choi, Y.M. and Lee, K.M. 2003. Practical Application of Life-Cycle Cost Effective Design and Rehabilitation of Civli Infrastructures. *3th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*. Lausanne, Switzerland.

Cho, H.N., Lee, K.M. and Choi, Y.M. 2004. Life-Cycle Cost Effective Optimum Design of Steel Bridges. *Journal of Constructional Steel Research* 60: 1585–1613.

Cho, H.N., Lee, K.M., Cha, C.J. and Kim, J.H. 2005. Life-Cycle Cost Effective Optimum Design of Orthotropic Steel Deck Bridges under Corrosion Environment, *4th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, FL, USA.

Cho, H.N., Lee, K.M., Choi, H.H. and Lim, J.K. 2006. Seismic reliability analysis of deteriorating bridges considering environmental stressor, *IABMAS. Proceedings of the 5th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems,* Seoul, Korea.

Cho, H.N. 2002. Life cycle cost effectiveness for design and rehabilitation of civil infrastructures. *Proceedings of SEWC02,* Yokohama, Japan.

Cho, H.N., Min, D.H. and Lee, K.M., 2001. Optimum Life-Cycle Cost Design of Orthotropic Steel Deck Bridges. *Steel Structures*. 1(2): 141–152.

Christian, J.T., Ladd, C.C. and Baecher, G.B. 1994. Reliability Applied to Slope Stability Analysis, *J. Geotech. Engrg.,* ASCE, Vol. 120, No. 12, pp. 2180–2207.

Christopher, B.R. and Di Maggio, J.A. 1984. Specifying Geotextiles, *Geotechnical Fabrics Report*, Vol. 2, No. 2, pp. 21–25.

Christopher, B.R. and Holtz, R.D. 1985. *Geotextile Engineering Manual*, U.S. Federal Highway Administration, Washington, D. C., FHWA-TS-86/203, 1044 pp.

Christopher, B.R. and Leshchinsky, D. 1991. Design of Geosynthetically Reinforced Slopes, State of the Art Paper, Session 8A on Earth Reinforcement, *Geotechnical Engineering Congress 1991*, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 998–1005.

Christopher, B.R., Gill, S.A., Giroud, J.P., Juran, I., Mitchell, J.K., Schiosser, F. and Dunnicliff, J. 1990. *Reinforced Soil Structures*, Vol. I: Design and Construction Guidelines, Federal Highway Administration, Report No. FHWA-RD-89-043, 287 pp.

Clough, R.W. and Penzien, J. 1993. *Dynamics of Structures*. 2nd Edition, McGraw-Hill, Inc.
Collin, J.G. Editor. 2001. *Design Manual for Segmental Retaining Walls*, Second Edition, National Concrete Masonry Association, NCMA Publication No. TR 127A, 289 pp.
Collins, K.R., Wen, Y.K. and Foutch, D.A. 1996. Dual-level design: a reliability-based methodology. *Earthquake Engineering and Structural Dynamics*, 25, pp.1433–1467.
Connor, R.J. and Fisher, J.W. 2000. In-service response of an orthotropic steel deck compared with design assumptions. *Transportation Research Record* 1696(1): 100–108.
Connor, R.J. and Fisher, J.W. 2001. Results of field measurements on the Williamsburg Bridge orthotropic deck: Final report on Phase III. *ATLSS Report No. 01-01*. Bethlehem, PA: ATLSS Engineering Research Center, Lehigh University.
Connor, R.J., Richards, S.O. and Fisher, J.W. 2003. Long-term monitoring of prototype orthotropic deck panels on the Bronx-Whitestone Bridge for fatigue evaluation. In K.M. Mahmoud (ed.), *2003 New York City Bridge Conference; Proc.,* New York City, October 20–21. Lisse: Swets & Zeitlinger.
Conti, M., Gregori, E. and Turi, G.A. 2005. Cross-Layer Optimization of Gnutella for Mobile Ad hoc Networks, *Proceedings, 6th ACM MobiHoc Symposium*, Urbana-Champaign, Illinois, USA, May 2005, pp. 343–354.
Cornell, C.A. 1968. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, **58**:1583–1606.
Cornell, C.A., Jalayer, F., Hamburger, R.O. and Foutch, D.A. 2002. Probabilistic basis for 2000 SAC FEMA steel moment frame guidelines. *Journal of Structural Engineering ASCE*, Vol. 128, No. 4, pp. 526–533.
Corotis, R.B. 2003. Socially relevant structural safety. *Applications of Statistics and Probability in Civil Engineering, Proc. ICASP 9* (Der Kiureghian, Madanat and Pestana, eds.), Millpress, Rotterdam, pp. 15–26.
Dedman, B. 2008a. Late Bridge Inspections Put Public at Risk, MSNBC Online available at http://www.msnbc.msn.com/id/20998261/ accessed 24 March.
Dedman, B. Ga. 2008b. Employee Faked Bridge Inspections, MSNBC Online available at http://www.msnbc.msn.com/id/23020686/ accessed 24 March.
DeGroot, D.J. and Baecher, G.B. 1993. Estimating Autocovariance of In-situ Soil Properties, *Journal of Geotechnical Engineering,* Vol. 119, No. 1, pp. 147–166.
Demers, C.E. and Fisher, J.W. 1990. A survey of localized cracking in steel bridge structures, Vol. 1. *Report No. FHWA-RD-89-166*. Washington D.C.: Federal Highway Administration, U.S. Department of Transportation.
Dexter, R.J., Connor, R.J. and Kaczinski, M.R. 1997. Fatigue design of modular bridge expansion joints. *NCHRP Report 402*. Washington, D.C.: Transportation Research Board.
Dong, X.J. The Seismic Performance Analysis of Electric Power Systems. *PhD Dissertation*, University of Southern California, June 2002.
Dowding, C.H. 1996. *Construction Vibrations*, Prentice Hall International Series in Civil Engieering and Engineering Mechanics, 379–39.
Duncan, J.M. 2000. Factors of Safety and Reliability in Geotechnical Engineering, *J. Geotech. and Geoenvir. Engrg.*, ASCE, Vol. 126, No. 4, pp. 307–316.
DuraCrete. 2000. *Modelling of Degradation*, Document BE95-1347/R4-5, Gouda.
Duxson, P., Provis, J.L., Lukey, G.C., Mallicoat, S.W., Kriven, W.M. and van Deventer, J.S.J. 2005. Understanding the relationship between geopolymer composition, microstructure and mechanical properties, *Colloids and Surfaces A – Physicochemical and Engineering Aspects*, Vol. 269, pp. 47–58.
Elias, V., Christopher, B.R. and Berg, R.R. 2001. *Mechanically Stabilized Earth Walls and Reinforced Soil Slopes–Design and Construction Guidelines*, U.S. Federal Highway Administration, National Highway Institute, Washington, D. C., Publication No. FHWA NHI-00-043, 394 pp.

Ellingwood, B.R. 1994. Probability-based codified design: past accomplishments and future challenges. *Structural Safety*, Vol. 13, No. 3, pp. 159–176.

Ellingwood, B.R. 1998. Reliability-based performance concept for building construction in *Structural Engineering Worldwide 1998*, Paper T161-4, Elsevier Science Ltd.

Ellingwood, B.R. 2007. Strategies for mitigating risk to buildings from abnormal load events. *International Journal of Risk Assessment and Management,* Vol. 7, Nos. 6/7, pp. 828–845.

Ellingwood, B. and Ang, A.H-S. 1972. *A Probabilistic Study of Safety Criteria for Design*, Structural Res. Ser. No. 387, University of Illinois.

Ellingwood, B. and Culver, C. 1977. Analysis of Live Loads in Office Buildings, *J. of Structural Div.*, ASCE, Vol. 103, No. ST8, pp. 1551–1560.

Ellingwood, B.R. and Galambos, T.V. 1982. Probability-Based Criteria for Structural Design. *Structural Safety*, Vol. 1, No. 1, pp. 15–26.

Ellingwood, E.R. and Naus, D.J. 1999. Condition Assessment and Maintenance of Aging Structure in Critical Facilities A probabilistic Approach, Case Study in Optimal Design and Maintenance Planning of Civil Infrastructure Systems, ASCE, pp 45–56.

Ellingwood, B.R., Celik, O.C. and Kinali, K. 2007. Fragility assessment of building structural systems in mid-America. *Earthquake Engineering & Structural Dynamics* Vol. 36, pp. 1935–1952.

Ellingwood, B.R., Galambos T.V., MacGregor J.G. and Cornell C.A. 1982a. Development of a probability based load criteria for American National Institute A58. NBS Special Publication 577.

Ellingwood, B.R., Galambos T.V., MacGregor J.G. and Cornell C.A. 1982b. Probability-based load criteria: load factors and load combinations. *Journal of Structural Division*, ASCE, 108(5), pp. 978–997.

Ellis, B.R. and Ji, T. 1997. Human-structure interaction in vertical vibrations. *Structures and Buildings*, *the Proceedings of Civil Engineers* 122:1, 1–9.

Elsener, B. 2002. Macrocell Corrosion of Steel in Concrete – Implications for Corrosion Monitoring, *J. of Cem. Concr. Compos.*, 24(1), 65–72.

Elsener, B. 2005. Corrosion Rate of Steel in Concrete – Measurements beyond the Tafel Law, *Corr. Sci.*, 47(12), 3019–3033.

Enright, M.P. and Frangopol, D.M. 1999a. Maintenance planning for deteriorating concrete bridges. *Journal of Structural Engineering*, ASCE, 125(12), 1407–1414.

Enright, M.P. and Frangopol, D.M. 1999b. Condition prediction of deteriorating concrete bridges using Bayesian updating. *Journal of Structural Engineering*, ASCE, 125(10), 1118–1124.

Enright, M.P. and Frangopol, D.M. 1999c. Reliability-based condition assessment of deteriorating concrete bridges considering load redistribution, *Structural Safety*, 21(2), 159–195.

Estes, A.C. and Frangopol, D.M. 2005. Life-cycle evaluation and condition assessment of structures. Chapter 36 in *Structural Engineering Handbook, 2nd Edition,* W-F. Chen and E.M. Lui, eds., CRC Press: 36–1 to 36–51.

EUROCODES. 2002. Basis of design and actions on structures, BS EN 1990:2002. British Standards Institution, London, United Kingdom.

Faber, M. and Stewart, M.G. 2003. Risk assessment for civil engineering facilities: a critical appraisal. *Reliability Engrg. & System Safety*, Vol. 80, pp. 173–184.

Fanjiang, G.-N., Ye, Q., Fernandez, O.N. and Taylor, L.R. 2004. Fatigue analysis and design of steel orthotropic deck for Bronx-Whitestone Bridge, New York City. *Transportation Research Record* 1892: 69–77.

Federal Emergency Management Agency (FEMA). 2000. *Recommended Seismic Design Criteria for New Steel Moment-Frame Buildings, FEMA 35.*

Federal Highway Administration. National Bridge Inventory, available online at http://www.fhwa.dot.gov/bridge/nbi.htm (accessed 22 May 2007).

Fenton, G.A. 1999. Estimating for Stochastic Soil Models, *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 125, No. 6, pp. 470–485.

Finno, R.J. and Gassman, S.L. 1998. Impulse Response Testing of Drilled Shafts, *J. Geotech. Geoenviron. Eng.*, Vol. 124, No. 10, pp. 965–975.

Fisher, J.W. 1984. *Fatigue and fracture in steel bridges: case studies*: John Wiley.

Fisher, J.W. and Pense, A.W. 1984. Mianus River Bridge material: Fracture surfaces and special tests. *Final Report*. Zetlin-Argo Structural Investigations, Inc. Connecticut Department of Transportation.

Fisher, J.W., Barthelemy, B.M., Mertz, D.R. and Edinger, J.A. 1980. Fatigue behavior of full-scale welded bridge attachments. *NCHRP Report 227*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Frank, K.H., Hirt, M.A. and McNamee, B.M. 1970. Effect of weldments on the fatigue strength of steel beams. *NCHRP Report 102*. Washington, D.C.: Highway Research Board.

Fisher, J.W., Hausammann, H., Sullivan, M.D. and Pense, A.W. 1979. Detection and repair of fatigue damage in welded highway bridges. *NCHRP Report 206*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Jin, J., Wagner, D.C. and Yen, B.T. 1990. Distortion-induced fatigue cracking in steel bridges. *NCHRP Report 336*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Mertz, D.R. and Zhong, A. 1983. Steel bridge members under variable amplitude long life fatigue loading. *NCHRP Report 267*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Nussbaumer, A., Keating, P.B. and Yen, B.T. 1993. Resistance of welded details under variable amplitude long-life fatigue loading. *NCHRP Report 354*. Washington, D.C.: Transportation Research Board.

Fisher, J.W., Yen, B.T. and Wang, D. 1990. Fatigue strength of riveted bridge members. *Journal of Structural Engineering* 116(11).

Fisher, J.W., Yen, B.T., Wang, D.W. and Mann, J.E. 1987. Fatigue and fracture evaluation for rating riveted bridges. *NCHRP Report 302*. Washington DC.: Transportation Research Board.

Frangopol, D.M. 1998. A Probabilistic Model Based on Eight Random Variables for Preventive Maintenance of Bridges, Progress meeting on optimum maintenance strategies for different bridge types, The Highways Agency, London.

Frangopol, D.M. 2003. Preventative Maintenance Strategies for Bridge Groups – Analysis, Final Project Report to the Highways Agency, London, U.K., March.

Frangopal, D.M. and Cheng, F.Y. (Eds.) 1996. *Advances in Structural Optimization*, ASCE.

Frangopol, D.M. and Furuta, H. (eds.). 2001. *Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, ASCE, Reston Virginia.

Frangopol, D.M. and Messervey, T.B. 2007. Risk assessment for bridge decision making, *Proceedings of the 4th Civil Engineering Conference in the Asia Region*, Taipei, Taiwan.

Frangopol, D.M. and Messervey, T.B. 2008a. Maintenance principles for civil structures using SHM. *Encyclopedia of Structural Health Monitoring*. Ed. C. Boller, F-K. Chang, and Y. Fujino. John Wiley & Sons, Ltd (Accepted for publication).

Frangopol, D.M. and Messervey, T.B. 2008b. Effect of SHM on reliability of bridges. Chapter 25 in *Monitoring Technologies for Bridge Management: State-of-the-Art*, A. Mufti and B. Bahkt, eds., Multi-Science Publishing Co., Ltd., England (invited paper, submitted).

Frangopol, D.M. and Liu, M. 2006. Life-cycle cost and performance of civil structures. *McGraw-Hill 2006 Yearbook of Science and Technology*, McGraw-Hill, New York: 183–185.

Frangopol, D.M. and Liu, M. 2007. Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost, *Structure and Infrastructure Engineering*; 3(1), 29–41.

Frangopol, D.M., Kong, J.S. and Gharaibeh, E.S. 2001. Reliability based life cycle management of highway bridges, *Journal of Computing in Civil Engineering,* ASCE 15(1), 27–34.

Frangopol, D.M., Lin, K.-Y., and Estes, A.C. 1997. Life-cycle cost design of deteriorating structures, *Journal of Structural Engineering*, ASCE, Vol. 123(10), pp.1390–1401.

Frankel, A. 1997. Simplified approach to incorporating uncertainty in the ground motion computation and mapping process. *ATC-35-2*, 8-1 to 8-38.

Frankel, S., Eydt, B., Owens, L. and Kent, K. 2006. *Guide to IEEE 802.11i: Establishing Robust Security Networks (Draft)*, National Institute of Standards and Technology (NIST), Gaithersburg, MD.

Frankel, A., Mueller, C., Barnhard, T., Perkins, D., Leyendecker, E.V., Dickman, N., Hanson, S. and Hopper, M. 1996. *National seismic-hazard maps: documentation*, *USGS Open-File Report* 96–532.

Friis-Hansen, P. 2005. Structuring of complex systems using Bayesian network. *Proceedings Workshop on Reliability Analysis of Complex Systems*, Technical University of Denmark, Lyngby, August 2004, pp. 111–133.

Fujino, Y. and Abe, M., 2004. Structural Health Monitoring – Current Status and Future. *Proceedings of the Second European Workshop on Structural Health Monitoring*, Munich, Germany.

Furuta, H. and Koyama, K. 2003. Optimal maintenance planning of bridge structures considering earthquake effects, *Proc. of IFIP TC7 Conference*, Antipolis, France.

Furuta, H., Frangopol, D.M. and Saito, M. 1998. Application of genetic algorithm to life-cycle cost maintenance of bridges, *Proc. of KKNN Symposium*, Seoul, Korea.

Furuta, H., Kameda, T. and Erami, M. 2005. A practical bridge management system using new multi-objective genetic algorithm, *Proc. of the 4th IABMAS Conference,* Porto, Portugal.

Furuta, H., Kameda, T. and Frangopol, D.M. 2004. Balance of structural performance measures, *Proc. of Structures Congress, Nashville, Tennessee, ASCE, May,* CD-ROM.

Furuta, H., Kameda, T., Fukuda, Y. and Frangopol, D.M. 2003a. Life-cycle cost analysis for infrastructure systems: life-cycle cost vs. safety level vs. service life, *Proc. of 3rd Life-Cycle Cost Analysis and Design of Infrastructure Systems*, Lausanne, Switzerland.

Furuta, H., Kameda, T., Nakahara, K. and Takahashi, Y. 2003b. Genetic algorithm for optimal maintenance planning of bridge structures, *Proc. of GECCO,* Chicago, US.

Galbraith, F.W. and Barton, M.V. 1970. Ground Loading from Footsteps. *Journal of Acoustical Society of America* 48, 1288–1292.

Gambino, S.J. and Gilbert, R.B. 1999. Modeling Spatial Variability in Pile Capacity for Reliability-Based Design, Analysis, Design, Construction and Testing of Deep Foundations, Roesset Edition, ASCE Geotechnical Special Publication No. 88, 135–149.

Gates, M. 1957. Empirical Formula for Predicting Pile Bearing Capacity, *Civil Engineering*, Vol. 27, No. 3, pp. 65–66.

Ge, J. 2006. On the Numerical Solution of Laplace's Equation with Nonlinear Boundary Condition for Corrosion of Steel in Concrete." *M.A.Sc. Thesis*, Carleton University, Ottawa, Canada.

Ghods, P., Isgor, O.B. and Pour-Ghaz, M. 2007. A Practical Method for Calculating the Corrosion Rate of Uniformly Depassivated Reinforcing Bars in Concrete, *Mater. Corros.*, 58(4), 265–272.

Ghosn, M. 2000. Development of Truck Weight Regulations Using Bridge Reliability Model, *Journal of Bridge Engineering,* 5(4): 293–303.

Ghosn, M. and Moses, F. 1998. Redundancy in highway bridge superstructures. NCHRP TRB Report 406, Washington, D.C., U.S.A.

Ghosn, M., Moses, F. and Wang, J. 2003. Design of Highway Bridges for Extreme Events, NCHRP TRB Report 489, Washington, D.C.

Gilbert, R.B. 2002. Questions for the Future of Risk Assessment in Environmental Geotechnics, *Proceedings*, 4th Intl. Congress on Environmental Geotechnics, Rio de Janeiro, Brazil, Vol. 2, 1007–1012.

Gilbert, R.B., Gambino, S.J. and Dupin, R.M. 1999. Reliability-Based Approach for Foundation Design without Site-Specific Soil Borings, *Proceedings*, OTC, Houston, Texas, OTC 10927, 631–640.

Gilbert, R.B., Najjar, S.S., Choi, Y.J. and Gambino, S.J. 2008. Practical Application of Reliability-Based Design in Decision Making, in *Reliability-Based Design in Geotechnical Engineering:* Computations and Applications, Phoon Edition, Taylor & Francis Books Ltd.

Gilbert, R.B., Tonon, F., Freire, J., Silva, C.T. and Maidment, D.R. 2006. Visualizing Uncertainty with Uncertainty Multiples. *Proceedings, GeoCongress 2006*, ASCE, Reston, Virginia.

Gilbert, R.B., Ward, E.G. and Wolford, A.J. 2001. A Comparative Risk Analysis of FPSO's with other Deepwater Production Systems in the Gulf of Mexico, *Proceedings*, OTC, Houston, Texas, OTC 13173.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. 1996. *Markov chain Monte Carlo in practice*. Chapman & Hall, New York, NY.

Gindy, M. and Nassif, H. 2006. Effect of bridge live load based on 10 years of WIM data. *Proceedings IABMAS'06*, Porto, Portugal.

Goldberg, D.E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Company, Inc.

Goldberg, D.E 1989. *Genetic Algorithms is Search Optimization and Machine Learning*, Addison – Wesley, Reading, Massachusetts.

Gordon, C.G. and Ungar, E.E. 1991. Generic Vibration Criteria for Vibration-Sensitive Equipment. *Proceedings of International Society for Optical Engineering (SPIE)*, San Jose, CA, 1619, 71–85.

Gransberg, D.D. 1996. Optimizing Haul Unit Size and Number Based on Loading Facility Characteristics, Journal of Construction Engineering and Management, ASCE, 123(3), pp 248–253.

GRL and Associates, Inc. 1995. *GRLWEAP Manual*, 4535 Emery Industrial Parkway, Cleveland, Ohio 44128, USA.

Gulikers, J. 2005. Theoretical Considerations on the Supposed Linear Relationship between Concrete Resistivity and Corrosion Rate of Steel Reinforcement, *Mater. Corros.*, 56(6), 393–403.

Haagensen, P.J., Statnikov, E.S. and Lopez-Martinez, L. 1998. Introductory fatigue tests on welded joint in high strength steel and aluminum improved by various methods including ultrasonic impact treatment (UIT). *IIW Doc. No. XIII-1748-98*. Paris: International Institute of Welding.

Haldar, A. and Martinez-Flores, R. 2008. A Novel Methodology for Assessment of Safety of Existing Structures. *Journal of Structural Engineering*, Vol. 35, No. 1, pp. 82–91.

Haldar, A., Martinez-Flores, R. and Katkhuda, H. 2008 (in press). Crack Detection in Existing Structures Using Noise-Contaminated Dynamic Responses. *Theoretical Applied Fracture Mechchanics*, doi:10.1016/j.tafmec.2008.04.007.

Haliburton, T.A., Anglin, C.C. and Lawmaster, J.D. 1978. *Selection of Geotechnical Fabrics for Embankment Reinforcement*, Report to US Army Engineer District, Mobil, Oklahoma State University, Stillwater, 138 pp.

Hall, J.F. (ed.) 1995. Northridge Earthquake of January 17, 1994: Reconnaissance Report. *Earthquake Spectra*, Supplement C to Vol. 11.

Hamburger, R.O., Foutch, D.A. and Cornell, C.A. 2003. Translating research to practice: FEMA/SAC performance-based design procedures. *Earthquake Spectra*, Vol. 19, No. 2, pp. 255–267.

Hannigan, P.J., Goble, G.G., Thendean, G., Likins, G.E. and Rausche, F. 1997. *Design and Construction of Driven Pile Foundations,* Workshop Manual, Vol. 1, *Publication No. FHWA-HI-97-014*, Federal Highway Administration, Washington, D.C.

Hashash, Y.M.A., Marulanda, C., Ghaboussi, J. and Jung, S. 2002. Systematic update of a deep excavation model using field performance data, *Computers and Geotechnics*.

Hassan, K.M. and O'Neill, M.W. 1998. *Structural Resistance Factors for Drilled Shafts with Minor Defects*, Final Rep., Phase I, Dept. of Civil and Environmental Engineering, Univ. of Houston, Houston.

Hiley, A. 1922. The Efficiency of the Hammer Blow and Its Effects with Reference to Piling, *Engineering*, Vol. 2, pp. 673.

Holtz, R.D. 1989. Treatment of Problem Foundations for Highway Embankments, *Synthesis of Highway Practice 147*, National Cooperative Highway Research Program, Transportation Research Board, 72 pp.

Holtz, R.D. 1990. Design and Construction of Geosynthetically Reinforced Embankments on Very Soft Soil, State of the Art Paper, Session 5, *Performance of Reinforced Soil Structures*, Proceedings of the International Reinforced Soil Conference, British Geotechnical Society, Glasgow, Scotland, A. McGown, K. Yeo, and K.Z. Andrawes, Editors, Th. Telford (London), pp. 391–402.

Holtz, R.D. 1995. Retaining Walls Reinforced with Geosynthetics: From Broms (1977, 1978) to the Present, *Proceedings* of the Bengt B. Broms Symposium on Geotechnical Engineering, Singapore, World Scientific, pp. 181–194.

Holtz, R.D., Christopher, B.R. and Berg, R.R. 1997. *Geosynthetic Engineering*, BiTech Publishers, Vancouver, British Columbia, Canada, 451 pp.

Holtz, R.D., Kramer, S.L., Hsieh, C.W., Huang, A.B. and Lee, W.F. 2001a. Failure Analysis of an 80 m High Geogrid Reinforced Wall in Taiwan, *Proceedings* of the 15th International Conference on Soil Mechanics and Geotechnical Engineering, Istanbul, Turkey, Volume 2, pp. 1159–1162.

Holtz, R.D., Shang, J.Q. and Bergado, D.T. 2001b. Soil Improvement, Chapter 15 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 429–462.

Hoshiya, M. and Saito, E. 1984. Structural Identification by Extended Kalman Filter. *Journal of Engineering Mechanics, ASCE*, Vol. 110, No. 12, pp. 1757–1770.

Hoshiya, M. and Sutoh, A. 1993. Kalman Filter- Finite Element Method in Identification. *Journal of Engineering Mechanics, ASCE*, V. 119, No. 2, pp. 197–210.

http://www.wecc.biz/main.html

Humphrey, D.N. and Rowe, R.K. 1991. Design of Reinforced Embankments—Recent Developments in the State-of-the-Art, State of the Art Paper, Session 8A on Earth Reinforcement, *Geotechnical Engineering Congress 1991*, Geotechnical Special Publication No. 27, ASCE, Vol. II, pp. 1006–1020.

Hunkeler, F. 2005. Corrosion in Reinforced Concrete: Processes and Mechanisms, Chapter 1, *Corrosion in Reinforced Concrete*, Edited by Hans Böhni, CRC Press, Washington DC.

ICC. 2000. *International Building Code*, International Code Council (ICC), USA.

Ioannou, P.G. and Martinez, J.C. 1996. Comparison of Construction Alternatives Using Matched Simulation Experiments, Journal of Construction Engineering and Management, ASCE, (112) 3, New York, NY.

Isgor, O.B. 2001. A Durability Model for Chloride and Carbonation Induced Steel Corrosion in Reinforced Concrete Members, *Ph.D. Thesis*, Carleton University, Ottawa, Canada.

Isgor, O.B. and Razaqpur, A.G. 2006. Modelling Steel Corrosion in Concrete Structures. *Mater. Struct.*, 39(287), 291–302.

Japan Society of Civil Engineers. *Report on Damage by Hanshin Awaji Earthquake*, 1996. (in Japanese).

Jazwinski, A.H. 1970. *Stochastic Process and Filtering Theory.* Academic Press.
Jensen, F.V. 2001. *Bayesian networks and decision graphs*. Springer, New York.
Jolliffe, I.T. 2002. *Principal Component Analysis*. 2nd ed. Springer Series in Statistics. Springer, New York, NY.
Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Transaction of the ASME*, pp. 35–45.
Kang, W.-H., Song, J. and Gardoni, P. 2007. Matrix-based system reliability method and applications to bridge networks. *Reliability Engineering & System Safety*, (accepted for publication).
Katkhuda, H. and Haldar, A. 2006. Defect Identification under Uncertain Blast Loading. *Optimization and Engineering Journal,* Vol. 7, No. 3, pp. 277–296.
Katkhuda, H. and Haldar, A. 2007. A Novel Health Assessment Technique with Minimum Information. *Structural Control and Health Monitoring*, published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/stc.221.
Katkhuda, H., Martinez-Flores, R. and Haldar, A., Health Assessment at Local Level With Unknown Input Excitation. *Journal of the Structural Engineering, ASCE,* Vol. 131, No. 6, pp. 956–965, June, 2005.
Keating, P. and Fisher, J.W. 1986. Evaluation of fatigue tests and design criteria on welded details. *NCHRP Report 286*. Washington, D.C.: Transportation Research Board.
Kern, F. 2002. Precise Determination of Volume with Terestical 3-D-Laserscanner. *Geodesy for Geotechnical and Structural Engineering II,* 531–534.
Khot, N.S., Berke, L. and Venkayya, V.B. 1979. Comparison of Optimality Criteria Algorithms for Minimum Weight Design of Structures, *AIAA J*., 17(2), 182–190.
Kinki Branch, Ministry of Land. Infrastructure and Transportation, *Road Traffic Census,* 1999. (in Japanese).
Kitano, H. (ed.) *Genetic Algorithm 3*, Tokyo, Sangyo-tosho, 1995. (in Japanese).
Klinzmann, C., Schnetgöke, R. and Hosser, D. 2006. A Framework for Reliability-Based System Assessment Based on Structural Health Monitoring, Proceedings of the 3rd European Workshop on Structural Health Monitoring, Granada, Spain.
Ko, Wen K. 1996. The future of sensor and actuator systems. *Sensors and Actuators*, V. 56, 193–197.
Koerner, R.M. 2005. *Designing with Geosynthetics*, 5th edition, Prentice-Hall, 706 pp.
Koerner, R.M. and Hsuan, Y.G. 2001. Geosynthetics: Characteristics and Testing, Chapter 7 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 173–196.
Koh, C.G. 1986. *Modelling of Dynamic Response of Elastomeric Isolation Bearings*, Ph.D. Thesis, University of California, Berkeley.
Koh, C.G., See, L.M. and Balendra, T. 1995. Damage Detection of Buildings: Numerical and Experimental Studies. *Journal of Structural Engineering, ASCE*, Vol. 121, No. 8, pp. 1155–60.
Kong, J.S. and Frangopol, D.M. 2004. Prediction of reliability and cost profiles of deteriorating structures under time- and performance-controlled maintenance. *Journal of Structural Engineering*, 130(12): 1865–1874.
Kong, J.S. and Frangopol, D.M., 2004. Cost-Reliability Interaction in Life-Cycle Cost Optimization of Deteriorating Structures, *Journal of Structural Engineering,* ASCE 130(11): 1704–1712.
Korea Infrastructure Safety and TEchnology Corporation. 2000. *Method for extension of service life of road bridges. Final report.* Ilsan, Korea Road and Transportation Association.
Korea Rail Network Authority. *Development of Life-Cycle Cost Analysis System of railroad bridges in design stage*. Daejeon, KRNA, 2007.
Korean Railroad. *Korean Railroad Integrated Facilities Information System*. DJ, Korean Railroad, 2003.

Kulhawy, F.H. 1996. From Casagrande's "Calculated Risk" to Reliability-based Design in Foundation Engineering, *J. Boston Society of Civil Engineers*, Vol. 11, No. 2, pp. 43–56,

Kumamoto, H. and Henley, E.J. 1996. Probabilistic Risk Assessment and Management for Engineers and Scientists, Second Edition, IEEE Press, New York.

Kunihiro, T., Inoue, K. and Fukuda, T. 1972. Atmospheric exposure of weathering steel. *Report Br. 71-08*. Tokyo Research Laboratory, Ministry of Construction.

Kwan, M.-P. and Lee, J. 2005. Emergency response after 9/11: the potential of real-time 3-D GIS for quick emergency response in micro-spatial environments, *Computing, Environment, and Urban Systems*, Vol. 29, pp. 93–113.

Ladd, C.C. 1991. Stability Evaluation During Staged Construction, 22nd Terzaghi Lecture, *Journal of Geotechnical Engineering*, ASCE, Vol. 117, No. 4, pp. 537–615.

Langseth, H. and L. Portinale. 2007. Bayesian networks in reliability. *Reliability Engineering & System Safety*, **92**:92–108.

Law, A.M. and Kelton, W.D. 2000. Simulation Modeling and Analysis, 3rd Edition. McGraw Hill, New York, NY, USA.

Lee, K.M. Cho, H.M and Cha, C.J. 2006. Life-cycle cost-effective optimum design of steel bridges considering environmental stressors, *Engineering Structure* 28(9): 1252–1265.

Lee, W.F. 2000. Internal Stability Analysis of Geosynthetic Reinforced Retaining Walls, Ph.D. dissertation, University of Washington, Seattle, 380 pp.

Lee, W.F., Holtz, R.D. and Allen, T.M. 1999. Full Scale Geosynthetic Reinforced Retaining.

Leeming, M.B., Mouchel, L.G., 1993. et al. The application of life cycle costing to bridges, Bridge management 2. Thomas Telford, London, 574–583.

Leemis, L.M. 1995. *Reliability: Probabilistic Models and Statistical Methods*, Prentice-Hall, Englewood Cliffs, N.J..

Leroueil, S. and Rowe, R.K. 2001. Embankments over Soft Soil and Peat, Chapter 16 in *Geotechnical and Geoenvironmental Handbook*, R.K. Rowe, Editor, Kluwer Academic Publishers, pp. 463–499.

Leung, K.W. and Papadimos, C.A. 1999. *Micro-Vibration Criteria for 300 mm and Beyond. Semiconductor Fabtech* 10th edition, Henley Publishing, London.

Li, D.Q., Zhang, L.M. and Tang, W.H. 2005. Evaluation of Reliability of Cross-hole Sonic Logging for Bored Piles, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 131, No. 9, pp. 1130–1138,

Liao, K.W., Wen, Y.K. and Foutch, D.A. 2007. Evaluation of 3D Steel Moment Frames under Earthquake Excitations. I Modeling *Journal of Structural Engineering*, ASCE, Vol 133, No. 3, pp. 462–471.

Liao, K.W., Wen, Y.K. and Foutch, D.A. 2007. Evaluation of 3D Steel Moment Frames under Earthquake Excitations. II Reliability and Redundancy *Journal of Structural Engineering*, ASCE, Vol 133, No. 3, pp. 471–480.

Liao, S.S.C., Druss, D.L., Neff, T.L. and Brenner, B.R. 1996. Just One More Borehole, and We'll Know for Sure. *Uncertainty in the Geologic Environment: From Theory to Practice.* Geotechnical Special Publication No. 58, Vol. 2, ASCE, New York, pp. 119–133.

Likins, G.E. 2004. Pile Testing-selection and Economy of Safety Factors, *Current Practices and Future Trends in Deep Foundations, Geotech. Spec. Publ. No. 125* (ed. by DiMaggio, J.A. and Hussein, M.H.), ASCE, Reston, Va., pp. 239–252.

Lim, J.K., Park, H.M., Park, K.H., Cho, H.N. and Kong, J.S. 2005. A Practical Life-Cycle Cost Analysis System for Steel Bridge Design and Maintenance Management Steel Bridges, *4th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems*, FL, USA.

Ling, X. and Haldar, A. 2004. Element Level System Identification with Unknown Input with Rayleigh Damping. *Journal of Engineering Mechanics, ASCE*, Vol. 130, No. 8, pp. 877–885.

Liu, M. and Frangopol, D.M. 2006a. Optimizing Bridge Network Maintenance Management under Uncertainty with Conflicting Criteria: Life Cycle Maintenance, Failure, and User Costs, *Journal of Structural Engineering,* ASCE, 131(11), 1835–1845.

Liu, M. and Frangopol, D.M. 2006b. Probability-Based Bridge Network Performance Evaluation, *Journal of Bridge Engineering*, ASCE, 11(5), 633–641.

Liu, T. and Weyers, R.W. 1998. Modelling the Corrosion Process in Chloride Contaminated Concrete Structures. *Cem. Concr. Res.*, 28(3), 365–379.

Lizotte, Y. and Bonates, E. 1987. Truck and Shovel Dispatching Rules Assessment Using Simulation, Mining Science and Technology, 5, pp 45–58.

Luco, N. and Cornell, C.A. Structure-specific scalar intensity measures for near source and ordinary earthquake ground motions. submitted to *Earthquake Spectr, 2002*.

Luping, T. 2002. *Calibration of the Electrochemical Methods for the Corrosion Rate Measurement of Steel in Concrete*, NORDTEST Project No. 1531-01, SP REPORT 2002:25, Swedish National Testing and Research Inst.

Lynch, J.P., Wang, Y., Lu, K-C., Hou, T-C. and Loh, C-H. 2006. Post-Seismic Damage Assessment of Steel Structures Instrumented with Self-Interrogating Wireless Sensors, *Proceedings of the 8th National Conference on Earthquake Engineering* (8NCEE), San Francisco, CA.

Maalej, M., Karasaridis, A., Hatzinakos, D. and Pantazopoulou, S.J. 1999. Spectral analysis of sensor data in civil engineering structure, *Computers and Structures*, V. 70, 675–689.

Malley, J.O. 2002. Update on US seismic design requirements for steel structures, *Proc. 7th US National Conference on Earthquake Engineering*, Boston, Mass. July 21–25.

Mark, A. Ehlen. 1997. Life-Cycle Costs of New Construction Materials. *Journal of Infrastructure systems*. 3(4): 129–133.

Marsh, P.S. and Frangopol, D.M. 2008. Reinforced concrete bridge deck reliability model incorporating temporal and spatial variations of probabilistic corrosion rate sensor data. Reliability Engineering & System Safety, Elsevier, 93(3), 394–409.

Martinez, J.C. 1996. STROBOSCOPE: State and Resource Based Simulation of Construction Processes. PhD Dissertation, University of Michigan, Ann Arbor, MI.

Martinez, J.C. 1998. Earthmover-Simulation Tool for Earthwork Planning, Proceedings of the 1998 Winter Simulation Conference, pg. 1263–1271, vol 2.

Martinez-Flores, R. and Haldar, A. 2007. Experimental Verification of a Structural Health Assessment Method without Excitation Information. *Journal Of Structural Engineering*, Vol. 34, No. 1, pp. 33–39, April-May.

Martinez-Flores, R., Haldar, A. and Katkhuda, H. 2006. Structural Health Assessment After an Impact. *Paper No. IMECE 2006-13718, ASME.*

Martinez-Flores, R., Katkhuda, H. and Haldar, A. 2008. A Novel Health Assessment Technique with Minimum Information: Verification. *International Journal of Performability Engineering,* Vol. 4, No. 2, pp. 121–140.

Maybeck, P. 1979. *Stochastic Models, Estimation, and Control*, Academic Press. U.K.

McGill, W.L., Ayyub, B.M. Kaminskiy, M. 2007. A Quantitative Asset-Level Risk Assessment and Management Framework for Critical Asset Protection, Risk Analysis International Journal, Society for Risk Analysis, 27(5), 2007, 1265–1281.

Mcvay, M.C., Kuo, C.L. and Singletary, W.A. 1998. *Calibrating Resistance Factors in the Load and Resistance Factor Design for Florida Foundations*, Final Rep., Dept. of Civil Engineering, Univ. of Florida, Gainesville.

McVay, M.C., Townsend, F.C. and Williams, R.C. 1992. Design of Socketed Drilled Shafts in Limestone, *J. Geotech. Eng.*, Vol. 118, No. 10, pp. 1626–1637.

Means. 1997. *Building Construction Cost Data/RS Means* (Annual) RS Means Company, Kingston, Massachusetts.

Mehrabian, A. and Haldar, A. 2002. Seismic Performance of Steel Frames with a Post-Northridge Connection, *Report No. CEEM-02-001*, Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, Arizona.
Melchers, R.E. 1987. Structural Reliability, Analysis and Prediction, Ellis Horwood Ltd., England.
Melchers, R.E. 1999. *Structural reliability: analysis and prediction*. John Wiley, Chichester, UK.
Messervey, T. 2007. Frangopol, D.M. A framework to incorporate structural health monitoring into reliability-based life-cycle bridge management models. *Proceedings CSMC5*, Rhodos, Greece.
Messervey, T.B. and Frangopol, D.M. 2008a. Probabilistic Treatment of Bridge Monitoring data and Associated Errors for Reliability Assessment and Prediction, *Proceedings, IALCCE08*, Como, Italy.
Messervey, T.B. and Frangopol, D.M. 2008b. Innovative Treatment of Monitoring Data for Reliability-Based Assessment, *Proceedings IABMAS'08,* Seoul, Korea.
Messervey, T.B. Frangopol, D.M. 2008c. Integration of Health Monitoring in Asset Management in a Life-Cycle Perspective, *Proceedings IABMAS'08,* Seoul, Korea.
Meyerhof, G.G. 1976. Bearing Capacity and Settlement of Pile Foundations, *J. Geotech. Engrg.*, Vol. 102, No. 3, pp. 195–228.
Mileti, D.S. 2007. Public Hazards Communication and Education: The State of the Art, www.incident.com/access/images/6/68.
Min, W. and Takada T. 2005. Macrospatial correlation model of seismic ground motions. *Earthquake Spectra*, **21**:1137–1156.
Ministry of Construction (MOC). 2003. *Technical Code for Testing of Building Foundation Piles (JGJ 106–2003),* Beijing, China.
Ministry of Land. 2002. Infrastructure and Transportation, *Design Specification of Highway Bridges*, Maruzen, (in Japanese).
Minnesota Department of Transportation Metro District. Fracture Critical Bridge Inspection, Bridge #9340 (Squirt Bridge), June 2006.
Miranda E. and Bertero, V.V. 1994. Evaluation of strength reduction factors for earthquake-resistant design. *Earthquake Spectra*, 10, pp. 357–379.
Mitchell, G. and Woodgate, R. 1971. *Floor Loadings in Office Buildings the Result of a Survey*, Current Paper 3/71, Building Research Station, Garston, Watford, England.
Modarres, M. 1993. What Every Engineer Should Know About Reliability and Analysis, Marcel Dekker, Inc., NY.
Modarres, M., Kaminskiy, M. Krivstov, V. 1999. Reliability Engineering and Risk Analysis: A Practical Guide, Marcel Decker Inc., New York, NY.
Mohraz, B., Hall, W. and Newmark, N. 1972. *A Study of Vertical and Horizontal Earthquake Spectra*, Div. of Reactor Standards, U.S. Atomic Energy Commission Washington, D.C.
Morinaga, S. 1990. Prediction of Service Life of Reinforced Concrete Buildings Based on the Corrosion Rate of Reinforcing Steel. *Proc. of the 5th International Conference on Durability of Building Materials and Components*, Brighton, U.K.
Moses, F., Schilling, C.G. and Raju, K.S. 1987. Fatigue evaluation procedures for steel bridges. *NCHRP Report 299*. Washington, D.C.: Transportation Research Board.
Munn, R.S. and Devereux, O.F. 1991. Numerical Modeling and Solution of Galvanic Corrosion Systems. Part I., *Corrosion*, 47(8), 612–618.
Nair, K.K. 2007. *Signal Processing Based Damage Detection Algorithms for Structural Health Monitoring*, Ph.D. Dissertation, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305.
Nair, K.K. and Kiremidjian, A.S. 2007. Time Series-Based Structural Damage Detection Algorithm Using Gaussian Mixtures Modeling, *Journal of Dynamic Systems, Measurement*, and Control, 129, 285–293.

Nair, K.K., Kiremidjian, A.S. and Law, K.H. 2006. Time Series-Based Damage Detection and Localization Algorithm with application to the ASCE Benchmark Structure, *Journal of Sound and Vibration*, 291(2), 349–368.

Nassar, A.A. and Krawinkler, H. 1992. Seismic design based on ductility and cumulative damage demands and capacities, *Nonlinear Seismic Analysis and Design of Reinforced Concrete Buildings*, ed. Fajfar, P. and Krawinkler, H., Elsevier Applied Sciences.

National Transportation Safety Board (NTSB). 1983. Highway Accident Report: Collapse of a suspended span of Interstate Route 95 Highway Bridge over the Mianus River, Greenwich, Connecticut. *Report No NTSB/HAR-84/03*.

NBS 577, *Development of a Probability Based Load Criterion for American National Standard A58*, National Bureau of Standards, 1980.

Neves, L.C. and Frangopol, D.M. 2004. Condition, safety, and cost profiles for deteriorating structures with emphasis on bridges, *Reliability Engineering and System Safety*, Vol. 89, pp. 185–198.

Neves, L.C., Frangopol, D.M. and Hogg, V. 2003. Condition-reliability-cost interaction in bridge maintenance, *Proceedings of the Ninth International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP9*, San Francisco, California, July 6–9, Millpress, Rotterdam, 2, 1117–1122.

Neves, L.A.C., Frangopol, D.M. and Petcherdchoo, A. 2006. Probabilistic lifetime-oriented multi-objective optimization of bridge maintenance: combination of maintenance types. Journal of Structural Engineering, ASCE, 132(11), 1821–1834.

Neville, A.M. 1995. *Properties of Concrete*, 4th Edition, Pearson Education Ltd., NY.

Nihon Sogo Research Institute. *Draft of Guideline for Evaluation of Investment on Road,* 1998. (in Japanese).

NIST. 2005. *Federal Building and Fire Safety Investigation of the World Trade Center Disaster: Final Report of the National Construction Safety Team on the Collapses of the World Trade Center Tower*, NIST, U.S.

Nodine, M.C., Wright, S.G., Gilbert, R.B., Ward, E.G. and Cheon, J.Y. 2008. Mudslides During Hurricane Ivan and an Assessment of the Potential for Future Mudslides in the Gulf of Mexico, Final Project Report, OTRC, Prepared for Minerals Management Service, Herndon, Virginia.

Noh, H., Nair, K.K., Kiremidjian, A.S. and Loh, C-H. 2007. Application of a Time Series-Based Damage Detection Algorithm to the Taiwanese Benchmark Experiment, *International Conference on Applications of Statistics and Probability in Civil Engineering*, CD Rom, Chiba, Japan. ISBN 978-0-415-45211-3.

Nojima, N. 2003. and Sugito, M. Development of a Probabilistic Assessment Model for Post-Earthquake Residual Capacity of Utility Lifeline Systems. J.E. Beavers, edition, *Advancing Mitigation Technologies and Disaster Response for Lifeline Systems*, Reston, Virginia: American Society of Civil Engineers, Technical Council on Lifeline Earthquake Engineering Monograph No.25, pp.707–716.

Northwood T.D. 1977. Isolation of Building Structures from Ground Vibration. *Isolation of Mechanical Vibration, Impact and Noise, AMD,* 1, 87–96.

Office of Construction and Facilities Management, *Data for design guide to maintenance and repair of infrastructure. Final report,* 2002.

O'Neill, M.W. and Reese, L.C. 1999. *Drilled Shafts: Construction Procedures and Design Methods*, *Publ. No. FHWA-IF-99-025*, Federal Highway Admin., Office of Implementation, McLean, Va.

O'Neill, M.W., Tabsh, S.W. and Sarhan, H.A. 2003. Response of Drilled Shafts with Minor Flaws to Axial and Lateral Loads, *Eng. Struc.*, Vol. 25, No. 1, pp. 47–56.

Orchant, C.J., Kulhawy, F.H. and Trautmann, C.H. 1988. Reliability-based Foundation Design for Transmission Line Structures, Vol. 2: Critical Evaluation of In Situ Test Methods, *EL-5507 Final report*, Electrical Power Institute, Palo Alto, Calif.

Oreta, A. and Tanabe, T. 1994. Element Identification of Member Properties of Framed Structures. *Journal of Structural Engineering, ASCE*, V. 20, No. 7, pp. 1961–1976.

Ove Arup and Partners Hong Kong Ltd. Construction of Barrettes and Diaphragm Walls at Union Square, Hong Kong, 2002.

Park, J., Bazzurro, P. and Baker, J.W. 2007. Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. in *Proceedings, 10th International Conference on Applications of Statistics and Probability in Civil Engineering*. Tokyo, Japan, Taylor & Francis Group, London, UK.

Park, Y.J., and Ang, A.H.S. 1985. Mechanistic seismic damage model for reinforced concrete, *Journal of Structural Engineering*, ASCE, Vol. 111, No. 4, 722–739.

Park, Y.J., Ang, A.H-S. and Wen, Y.K. 1985. Seismic damage analysis of reinforced concrete builidngs. *Journal of Structural Engineering*, ASCE, 111(4), pp.740–757.

Pate-Cornell, E. 1994. Quantitative safety goals for risk management of industrial facilities. *Structural Safety*, Vol. 13, No. 3, pp. 145–157.

PDCA. *Recommended Design Specifications for Driven Bearing Piles*, Pile Driving Contractors Association (PDCA), Boulder, CO, USA, 2001.

Pearl, J. 1988. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers, San Francisco, CA.

Peck, R. 1977. Pitfalls of Overconservatism in Geotechnical Engineering, *Civil Engineering*, ASCE, pp. 62–66.

Peil, U. 2003. Life-Cycle Prolongation of Civil Engineering Structures via Monitoring, *Proceedings of the 4th International Workshop on Structural Health Monitoring*, Stanford, CA.

Piringer, S. 1993. Whole-Life Costing of Steel Bridges, Bridge Management, Thomas Telford, London, U.K., 584–593.

PONTIS. user manual and product web site, Cambridge Systematics, available through the AASHTO products office at http://www.camsys.com/tp_inframan_pontis.htm (accessed 2 Nov 2007).

Pour-Ghaz, M. 2007. A Novel Approach for Practical Modelling of Steel Corrosion in Concrete, *M.A.Sc Thesis*, Carleton University, Ottawa, Canada.

Prieto, R. 2002. Keynote Presentation, FIATECH, *Homeland Security and Roadmap Implementation Workshop*, Nov. 13–15, National Conference Center, Lansdowne, VA.

Rafiq, M.I. 2005. Health Monitoring in Proactive Reliability Management of Deteriorating Concrete Bridges, University of Surrey, UK.

Raupach, M. and Gulikers, J. 1999. A Simplified Method to Estimate Corrosion Rates – A New Approach Based on Investigations of Macrocells, *8th International Conference on Durability of Building Materials and Components*, Vancouver, Canada.

Rausche, F., Thendean, G., Abou-Matar, H., Likins, G.E. and Goble, G.G. 1996. Determination of Pile Driveability and Capacity from Penetration Tests, *FHWA Report DTFH61-91-C-00047*, Federal Highway Administration, Washington, D.C.

Reid, S.G. 2007. Acceptable risk criteria. *Progress in Structural Engineering and Materials* Vol. 2, No. 4, pp. 254–262.

Rice, J.A. 1999. *Mathematical Statistics and Data Analysis, Second Edition*, Duxbury Press, Second Edition, New York.

RILEM TC 154-EMC. Half-cell Potential Measurements – Potential Mapping on Reinforced Concrete Structures, *Mater. Struct.*, 36(7), 461–471, 2003.

Roberts, J.E. and Shepard, R. 2002. Bridge Management for the 21st Century. *Proceedings, IABMAS'02*, Barcelona, Spain.

Romano, J.J. 1997. Modeling Bridge Deck Maintenance Costs, unpublished Master's thesis, Dept. of civil and Environmental Engineering, Lehigh University, Bethlehem, PA.

Romstad, K.M. and Wang, C.K. 1968. Optimum Design of Framed Structures, Journal of Structural Division, ASCE, Vol. 94, No. ST12.

Romstad, K.M., Hutchinson, J.R. and Runge, K.H. 1972. Design Parameter Variation and Structural Response, International Journal for Numerical Methods in Engineering, Vol. 5, Issue 3, pp. 337–349.

Rose, A., Benavides, J., Chang, S.E., Szczesniak, P. and Lim, D. 1997. The Regional Economic Impact of an Earthquake: Direct and Indirect Effects of Electricity Lifeline Disruptions. *Journal of Regional Science*, Vol. 37, No. 3, pp. 437–458.

Roy, S. and Fisher, J.W. 2005. Enhancing fatigue strength by ultrasonic impact treatment. *International Journal of Steel Structures* 5(3): 241–252.

Roy, S. and Fisher, J.W. 2006. Modified AASHTO design S-N curves for post-weld treated welded details. *Journal of Bridge Structures – Assessment, Design and Construction* 2(4): 207–222.

Roy, S., Fisher, J.W. and Yen, B.T. 2003. Fatigue resistance of welded details enhanced by ultrasonic impact treatment (UIT). *International Journal of Fatigue* 25(9–11): 1239–1247.

SAA. 1995. Australian Standards: *Piling – Design and Installation AS-2159*, Standards Association of Australia (SAA), Homebush, NSW, Australia.

Saadeghvaziri, M.A. and Feng, M. 2001. Experimental and Analytical Study of Base-Isolation for Electric Power Equipment. *Research Progress and Accomplishments 2000-2001*, Multidisciplinary Center for Earthquake Engineering Research, May, pp. 29–40.

SAC Steel Project. *Seismic Design Criteria for Steel Moment-Frame Structures, SAC CD01-, 2000*.

Sarhan, H.A., O'Neill, M.W. and Tabsh, S.W. 2004. Structural Capacity Reduction for Drilled Shafts with Minor Flaws, *ACI Struct. J.,* Vol. 101, No. 3, pp. 291–297.

Sarma, K.C. and Adeli, H. 2000. Fuzzy Genetic Algorithm for Optimization of Steel Structures, *Journal of Structural Engineering, ASCE*, 126(5), 596–604.

Schexnayder, C., Weber, S.L. and Brooks, B.T. 1999. Effect of Truck Payload Weight on Production, Journal of Construction Engineering and Management, ASCE, 125(1), pp 1–7.

Schilling, C.G., Klippstein, K.H., Barsom, J.M. and Blake, G.T. 1978. Fatigue of welded steel bridge members under variable amplitude loading. *NCHRP Report 188*. Washington, D.C.: Transportation Research Board.

Schulz, Whitten L., Joel P. Conte and Eric Udd. 2001. Long Gage Fiber Optic Bragg Grating Strain Sensors to Monitor Civil Structures, *Proceeding of SPIE: Smart Structures & Material 2001: Smart Systems for Bridges, Structures, and Highway*, V. 4330, 56–65.

Senior Seismic Hazard Analysis Committee (SSHAC) *Probabilistic seismic hazard analysis: a consensus methodology, 1995*.

Shastry, C.R., Friel, J.T. and Townsend, H.R. 1987. Sixteen-year atmospheric corrosion performance of weathering steels in marine, rural and industrial environments. In S.W. Dean and T.S. Lee (eds.), *Symposium on Degradation of Metals in the Atmosphere, Proc., ASTM STP 965*. Philadelphia, PA: ASTM International.

Shinozuka, M. and Chang, S.E. 2004. Evaluating the Disaster Resilience of Power Networks and Grids. *Modeling Spatial Economic Impacts of Disasters*, Springer-Verlag Edited by Y. Okuyama and S.E. Chang.

Shinozuka, M. and Dong, X. 2002. Seismic Performance Criteria for Lifeline Systems. *Proceedings of the Eighth U.S. – Japan Workshop on Earthquake Resistant Design of Lifeline Facilities and Countermeasures Against Soil Liquefaction*, Tokyo, Japan, Dec. 15–18.

Shinozuka, M., Feng, M.Q., Lee, J. and Naganuma, T. 2000. Statistical analysis of fragility curves. *Journal of Engineering Mechanics, ASCE*, Vol. 126, No. 12, pp. 1224–1231.

Shinozuka, M., Cheng, T.C., Jin, X., Dong, X. and Penn, D., 2002. System Performance Analysis of Power Networks. *Proceedings of the Seventh U.S. National Conference on Earthquake Engineering* (7NCEE), Boston, Massachusetts, July 21–25.

Shoji, M., Fujino, Y. and Abe, M. Optimization of seismic damage allocation of viaduct systems, *Proc. of JSCE*, No. 563, 79–94, 1997. (in Japanese)

Shome, N., Cornell, A.C., Bazzurro, P. and Carballo, J.E. 1998. Earthquakes, records, and nonlinear responses. *Earthquake Spectra*, 14(3), pp. 469–500, 1998.

Singhal, A. and Kiremidjian, A.S. Method for probabilistic evaluation of seismic structural damage. *Journal of Structural Engineering, ASCE*, Vol. 122, No. 12, pp. 1459–1467, 1996.

Sivakumar, B., Khoury, E., Fisher, J.W. and Wright, W.J. 2003. Hoan bridge brittle fracture and retrofit design. In K.M. Mahmoud (ed.), *2003 New York City Bridge Conference; Proc.,* New York City, October 20–21. Lisse: Swets & Zeitlinger.

Sohn, H., Farrar, C.R., Hunter, H.F. and Worden, K. 2001. Applying the LANL Statistical Pattern Recognition Paradigm for Structural Health Monitoring to Data from a Surface-Effect Fast Patrol Boat, *Los Alamos National Laboratory Report LA-13761-MS*, Los Alamos National Laboratory, Los Alamos, NM 87545.

Somerville, P.G., Smith, N., Punyamurthula, S., and Sun, J. 1997. Development of ground motion time histories for Phase 2 of the FEMA/SAC Steel Project. *Report No. SAC/BD-97/04, SAC Joint Venture*, Sacramento, California.

Statnikov, E.S. 1997a. Applications of operational ultrasonic impact treatment (UIT) technologies in production of welded joint. *IIW Doc. No. XII-1667-97.* Paris: International Institute of Welding.

Statnikov, E.S. 1997b. Comparison of post-weld deformation methods for increase in fatigue strength of welded joints. *IIW Doc. No. XII-1668-97.* Paris: International Institute of Welding.

Stern, M. and Geary, A.L. 1957. Electrochemical Polarization, I. A Theoretical Analysis of the Shape of Polarization Curves, *J. Electrochem. Soc.*, 104(12), 56–63.

Stewart M.G. and Hossain, M.B. 2001. Time-dependant Deflection, Serviceability Reliability and Expected Cost for RC beams, Structural Safety and Reliability, Corotis et. al. (eds).

Stewart, M.G. Melchers, R.E. 1997. *Probabilistic Risk Assessment of Engineering Systems*, Chapman & Hall, London.

Straser, E.G. and Kiremidjian, A.S. 1998. Modular Wireless Damage Monitoring System for Structures, *Technical Report No. 128*, John A. Blume Earthquake Engineering Center, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA.

Structural Engineering Association of California (SEAOC). *Vision 2000, Performance-based Seismic Engineering of Buildings, 1995*.

Takamori, H. and Fisher, J.W. 2000. Tests of large girders treated to enhance fatigue strength. *Transportation Research Record* 1696: 93–99.

Tan, S.L. 1990. Stress-Deflection Characteristics of Soil Soils Overlain with Geosynthetics, MSCE Thesis, University of Washington, Seattle, 146 pp.

Tanaka, S., Shinozuka, M., Schiff, A. and Kawata, Y. 1997. Lifeline Seismic Performance of Electric Power Systems during the Northridge Earthquake. *Proceedings of the Northridge Earthquake Research Conference*, Los Angeles, California, August 20–22.

Tang, W.H. 1984. Principles of Probabilistic Characterization of Soil Properties, *Probabilistic Characterization of Soil Properties: Bridge between Theory and Practice*, D.S. David and H.-Y. Ko (Eds.), ASCE, New York, pp. 74–89.

Tang, W.H. 1989. Uncertainties in Offshore Axial Pile Capacity, *Geotechnical Special Publication*, No. 27, Vol. 2, ASCE, New York, pp. 833–847.

Tang, W.H. 1987. Updating Anomaly Statistics – Single Anomaly Case, *Structural Safety*, No. 4, pp. 151–163.

Tang, W.H. and Gilbert, R.B. 1993. Case Study of Offshore Pile System Reliability. *Proc., 25th Offshore Technology Conference*, Houston, TX, pp. 677–686.

Tang, W.H., Sidi, I. and Gilbert, R.B. 1989. Average Property in a Random Two-state Medium, *Journal of Engineering Mechanics*, ASCE, Vol. 115, No. 1, pp. 131–144.

Tang, W.H., Zhang, L.M. and Zheng, Y.R. Dealing with Uncertainty in Engineering Design for Large-scale Gravel Soil Slopes in the Three Gorges Reservoir Zone, Keynote lecture paper, *Geohazards – Technical, Economical and Social Risk Evaluation,* 18–21 June 2006, Lillehammer, Norway. Farrokh Nadim (ed.), Berkeley Electronic Press, in CD Rom.

Terzaghi, K., Peck, R.B. and Mesri, G. 1996. *Soil Mechanics in Engineering Practice*, 3rd Edition, Wiley, 549 pp.

Thasnanipan, N., Maung, A.W., Navaneethan, T. and Aye, Z.Z. 2000. Non-Destructive Integrity Testing on Piles Founded in Bangkok Subsoil, *Proc., 6th Int. Conf. on the Application of Stress-wave Theory to Piles*, S. Niyama and J. Beim, eds., A.A. Balkema, Rotterdam, pp. 171–177.

Thomas, S., Onley, R.E. and Morich, R.S. 1999. Design of a Crash Survivable Locomotive Event Recorder, *International Symposium on Transportation Recorders*, NTSB, pp. 313–323.

Thompson, M.H. 1999. A Vision of Future Crash Survivable Recording Systems, *International Symposium on Transportation Recorders*, NTSB, pp. 337–349.

Tibshirani, R., Walther, G. and Hastie, T. 2001. Estimating the Number of Clusters in a Dataset via the Gap Statistic, *Journal of the Royal Statistical Society*: Series B, 63(2), pp. 411–423.

Tijou, J.C. 1984. Integrity and Dynamic Testing of Deep Foundations – Recent Experiences in Hong Kong (1981–83), *Hong Kong Engineer*, Vol. 12, No. 9, pp. 15–22.

Tilly, G.P. and Nunn, D.E. 1980. Variable amplitude fatigue in relation to highway bridges. *Proc. the Institution of Mechanical Engineers (London)* 194: 259–267.

Tobias, D.H. and Foutch, D.A. 1997. Reliability Based Method for Fatigue Evaluation of Railway Bridges, *Journal of Bridge Engineering*, ASCE, V. 2, No. 2, 53–60.

Truman, K.Z. and Cheng, F.Y. 1990. Optimum Assessment of Irregular 3-D Seismic Buildings, *Journal of Structural Engineering*, ASCE, Vol.116, No.12, pp. 3324–3337.

Tufte, E.R. 1990. Envisioning Information, Graphics Press, Cheshire, Connecticut.

Turner, M.J. 1997. *Integrity Testing in Piling Practice*, Construction Industry Research and Information Association, CIRIA Report 144, London.

U.S.-Canada Power System Outage Task Force. 2003. Interim Report: Causes of the August 14th Blackout in the United States and Canada, ftp://www.nerc.com/pub/sys/all_updl/docs/blackout/814BlackoutReport.pdf

Uhlig, H.H. and Revie, R.W. 1985. *Corrosion and Corrosion Control*, 3rd Edition, John Wiley & Sons, New York.

Ungar, E.E. and White R.W., 1979. Footfall Induced Vibration of Floors Supporting Sensitive Equipments. *Sound and Vibration,* 10.

USACE, 2006. Interagency Performance Evaluation Task Force Draft Report on Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System, Draft Volume VIII – Engineering and Operational Risk and Reliability Analysis, USACE, Washington, DC. https://IPET.wes.army.mil

Vamvatsikos, D. and Cornell, C.A. 2002. Incremental dynamic analysis. *Journal of Earthquake Engineering and Structural Dynamics*, 31(3), pp. 491–514.

Vanmarcke, E. 2008. *Random Fields*, 2nd Edition, World Scientific Publishing.

Vanmarcke, E. 1974. Decision Analysis in Dam Safety Monitoring, *Proc. Eng. Found. Conf. on the Safety of Dams*, Henniker, New Hampshire, ASCE, pp. 127–148.

Venkayya, V.B. and Cheng, F.Y. 1976, Resizing of Frames Subjected to Ground Motion, *Proc. of the International Symposium on Earthquake Structural Engineering*, University of Missouri-Rolla, Vol. 1, pp. 597–612.

Venkayya, V.B., Khot, N.S. and Reddy, V.S. 1969. *Energy Distribution in an Optimum Structural Design*, Technical Report AFFDL-TR_68-156, Wright-Patterson Air Force Base, Ohio.

Vo, P.H. and Haldar, A. 2008a. Health Assessment of Beams – Theoretical Formulation and Analytical Verification. *Structure and Infrastructure Engineering*, Vol. 4, No. 1, pp. 33–44.

Vo, P.H. and Haldar, A. 2008b. Health Assessment of Beams – Experimental Verification. *Structure and Infrastructure Engineering*, Vol. 4, No. 1, pp. 45–56.

Vo, P.H. and Haldar, A. 2003. Post Processing of Linear Accelerometer Data in System Identification. *Journal of Structural Engineering*, Vol. 30, No. 2, pp. 123–130.

Walls, J. and Smith, M.R. 1998. *Life-Cycle Cost Analysis In Pavement Design – Interim Technical Bulletin*. Fhwa-Sa-98-079. Research Report. Washington, DC: Federal Highway Administration.

Walls: A Numerical Parametric Study, *Proceedings* of Geosynthetics '99, Boston, Massachusetts, Industrial Fabrics Association International, Vol. 2, pp. 935–948.

Wang, C.-H. and Wen, Y.K. 2000. Evaluation of pre-Northridge low-rise steel buildings-Part I, Modeling. *Journal of Structural Engineering*, *ASCE*, 126(10), pp. 1151–1160.

Wang, C.-H. and Wen, Y.K. 2000. Evaluation of pre-Northridge low-rise steel buildings-Part II, Reliability. *Journal of Structural Engineering*, ASCE, 126(10), pp. 1161–1169.

Wang, D. and Haldar, A. 1994. An Element Level SI with Unknown Input Information. *Journal of Engineering Mechanics Division, ASCE*, Vol. 120, No. 1, pp. 159–176.

Wang, D. and Haldar, A. System Identification with Limited Observations and Without Input. *J. of Engineering Mechanics, ASCE*, 123(5) 1997. 504–511.

Wang, Y., Lynch, J.P. and Law, K.H. 2006. Wireless sensing, actuation and control – with applications to civil structures, Intelligent Computing in Engineering and Architecture (Ed. Smith, I.F.C.), Springer, Berlin, Germany.

Watanabe, E., Frangopol, D.M. and Utsunomiya, T. 2004. eds., *Bridge Maintenance, Safety, Management and Cost*, A.A. Balkema, Swets & Zeitlinger B.V., Lisse, The Netherlands.

Watkins, Steve E., Unser, John F. Antonio Nanni, Chandrashekhara, K. and Abdeldjelil Belarbi. 2001. Instrumentation and manufacture of a smart composite bridge for short-span applications. *Proceeding of SPIE: Smart Structures & Material 2001: Smart Systems for Bridges, Structures, and Highway*, V. 4330, 147–156.

Webb, G.R., Tierney, K., and Dahlhamer, J.M. 2000. Businesses and Disasters: Empirical Patters and Unanswered Questions. *Natural Hazards Reviews*, Vol. 1, No. 2, pp. 83–90.

Weber, P. and L. Jouffe. 2006. Complex system reliability modeling with Dynamic Object Oriented Bayesian Networks (DOOBN). *Reliability Engineering & System Safety*, **91**:149–162.

Wen, Y.K. 1989. Methods of random vibration for inelastic structures. *Applied Mechanics Reviews*, ASME, 42(2), pp. 39–52.

Wen, Y.K. 1990. *Structural Load Modeling and Combination for Performance and Safety Evaluation*, Elsevier, Development in Civil Engineering Series, Vol.31.

Wen, Y.K. and Ellingwood, B.R. 2005. The role of fragility assessment in consequence-based engineering. *Earthquake Spectra* EERI, Vol. 21, No. 3, pp. 861–877.

Wen, Y.K. and Foutch, D.A. 1997. Proposed statistical and reliability framework for comparing and evaluating predictive models for evaluation and design and critical issues in developing such framework. *Report No. SAC/BD-97/03, SAC Joint Venture*, Sacramento, California.

Wen Y.K. and Kang Y.J. 2000. Minimum lifecycle cost structural design against natural hazards, Structural Research Series No. 629, University of Illinois at Urbana-Champaign.

Wen, Y.K. and Song, S.-H. 2003. Structural Reliability/Redundancy under Earthquakes, *Journal of Structural Engineering*, Vol. 129, Issue 1, pp. 56–67.

Wen, Y.K. and Wu, C.L. 2001. Uniform hazard ground motions for mid-America cities. *Earthquake Spectra*, 7(2), pp. 359–384.

Wilson, A.G., L.A. McNamara and Wilson, G.D. 2007. Information integration for complex systems. *Reliability Engineering & System Safety*, **92**:121–130.

Working Group on California Earthquake Probabilities. Seismic hazards in southern California; probable earthquakes, 1994 to 2024. *Bulletin Seismological Society of America*, 85, pp. 379–439, 1995.

Works Bureau. 2000. *Enhanced Quality Supervision and Assurance for Pile Foundation Works,* Works Bureau Technical Circular No. 22/2000, Hong Kong.

Wright, W.J. 1996. Post-weld treatment of a welded bridge girder by ultrasonic hammer peening. *FHWA Internal Research Report.* McLean, VA: Federal Highway Administration, Turner Fairbank Highway Research Center.

Wright, W.J., Fisher, J.W. and Kaufmann, E.J. 2003. Failure analysis of the Hoan Bridge fractures. In K.M. Mahmoud (ed.), *2003 New York City Bridge Conference; Proc.,* New York City, October 20–21. Lisse: Swets & Zeitlinger.

Wu, T.H., Tang, W.H., Sangrey, D.A. and Baecher, G.B. 1989. Reliability of Offshore Foundations-State of the Art, *J. Geotech. Engrg. Div.*, ASCE, Vol. 115, No. 2, pp. 157–178.

Yamada, K. and Kikuchi, Y. 1984. Fatigue tests of weathered welded joints. *ASCE Journal of the Structural Division* 110(9).

Yi, S.C., Cho, H.N., Hwang, Y.K. and Lee, K.M. 2003. Practical Life-Cycle Cost Effective Optimum Design of Steel Bridges. *3th International Workshop on Life-Cycle Cost Analysis and Design of Civil Infrastructure Systems.* Lausanne, Switzerland.

Zhang, L.M. 2004. Reliability Verification Using Proof Pile Load Tests, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 130, No. 11, pp. 1203–1213.

Zhang, L.M. and Ng, A.M.Y. 2005. Probabilistic Limiting Tolerable Displacements for Serviceability Limit State Design of Foundations, *Geotechnique*, Vol. 55, No. 2, pp. 151–161.

Zhang, L.M., Li, D.Q and Tang, W.H. 2005. Reliability of Bored Pile Foundations Considering Bias in Failure Criteria, *Canadian Geotech. Journal*, Vol. 42, No. 4, pp. 1086–1093.

Zhang, L.M., Li, D.Q. and Tang, W.H. 2006a. Impact of Routine Quality Assurance on Reliability of Bored Piles, *J. Geotech. Geoenviron. Engrg.*, ASCE, Vol. 132, No. 5, pp. 622–630.

Zhang, L.M., Li, D.Q. and Tang, W.H. 2006b. Level of Construction Control and Safety of Driven Piles, *Soils and Foundations*, Vol. 46, No. 4, pp. 415–426.

Zhang, L.M. Shek, M.P., Pang, W.H. and Pang, C.F. Knowledge-based Pile Design using a Comprehensive Database. *Geotech. Engrg., Proc. of the Institution of Civil Engineers*, Vol. 159, No. GE3, pp. 177–185, 2006c.

Zhang, L.M., Tang, W.H. and Ng, C.W.W. 2001. Reliability of Axially Loaded Driven Pile Groups, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 127, No. 12, pp. 1051–1060.

Zhang, L.M., Tang, W.H., Zhang, L.L. and Zheng, J.G. 2004. Reducing Uncertainty of Prediction from Empirical Correlations, *J. Geotech. Geoenviron. Engrg.,* ASCE, Vol. 130, No. 5, pp. 526–534.

Zhao, Z. and Haldar, A. 1996. Fatigue Damage Evaluation and Updating Using Nondestructive Inspections. *Journal of Engineering Fracture Mechanics*, Vol. 53, No. 5, pp. 775–788.

Zhao, Z., Haldar, A. and Breen, F.L. 1994. Fatigue Reliability Updating Through Inspections for Bridges. *Journal of the Structural Engineering, ASCE*, Vol. 20, No. 5, pp. 1624–1642.

Zimmerman, K.A. 1997. Guidelines for Using Economic Factors and Maintenance Costs in Life-Cycle Cost Analysis, Technical Report No. SD96-08-F (NTIS PB98-110638), Applied Pavement Technology, Inc.

Author Index

About the Corresponding Authors

Ayubb, Bilal M. – Dr. Ayyub is Professor of Civil and Environmental Engineering at the University of Maryland, College Park and Director of the Center for Technology and Systems Management. He is also a researcher and consultant in the areas of risk analysis, uncertainty modeling, decision analysis, and systems engineering. He completed his B.S. degree in civil engineering in 1980 from the University of Kuwait, and completed both the M.S. (1981) and Ph.D. (1983) degrees in civil engineering at the Georgia Institute of Technology. Dr. Ayyub is a fellow of ASCE, ASME and SNAME, and has served the engineering community in various capacities through societies that include ASNE, ASCE, ASME, SNAME, IEEE-CS, and NAFIPS. He is the author and co-author of more than 500 publications in journals and conference proceedings, and reports. His publications include several books, including *Uncertainty Modeling and Analysis for Engineers and Scientists*, 2006; *Risk Analysis in Engineering and Economics*, 2003; and *Uncertainty Modeling and Analysis for Engineers and Scientists*, 2006. He is a multiple recipient of the ASNE "Jimmie" Hamilton Award for the best papers in the Naval Engineers Journal in 1985, 1992, 2000 and 2003. From ASCE the "Outstanding Research Oriented Paper" in the Journal of Water Resources Planning and Management in 1987, the Edmund Friedman Award in 1989, the Walter Huber Research Prize in 1997, and the K. S. Fu Award of NAFIPS in 1995.

Cheng, Franklin – After receiving the BS degree (1960) from the National Cheng-Kung University in Taiwan, and the MS degree (1962) from the University of Illinois at Urbana-Champaign, Dr. Cheng gained industrial experience with C.F. Murphy and Sargent & Lundy in Chicago, Illinois. He then received a PhD degree (1966) in civil engineering from the University of Wisconsin-Madison. He joined the University of Missouri-Rolla (now Missouri University of Science and Technology) in 1966 and was appointed Curators' Professor in 1987, the highest professorial position in the University system, and Emeritus Professor since 2000. In 2007, the American Society of Civil Engineers recognized Dr. Cheng's accomplishments by electing him to Honorary Membership (currently called Distinguished Membership). Dr. Cheng received four honorary professorships from China and chaired seven of his 24 National Science Foundation delegations to various countries for research cooperation. He is the author of more than 280 publications, including three text books, *Matrix Analysis of Structural Dynamics – Applications and Earthquake Engineering, Dynamic Structural Analysis,* and *Smart Structures - Innovative Systems for Seismic Response Control*, and a forthcoming book entitled *Structural Optimization – Application of Dynamic Loading.*

Cho, Hyo-Nam – Dr. Cho is currently Professor Emeritus at Hanyang University, Ansan, Korea, Chairman of UNICONS Co. Ltd. a consulting firm, and Editor-in-chief of IJOSS (International Journal of Steel Structure). He is also Vice President of IALCCE (International Association of Life Cycle Civil Engineering). Dr. Cho obtained his BS degree from the Korean Military Academy in 1967, and both the MS and PhD degrees in Structural Engineering from Michigan State University in 1970 and 1972, respectively. In 1973, he joined the Korean Military Academy, as Professor in the Department of Civil Engineering and held various academic positions until 1988. Since 1988, Dr. Cho has been Professor in the Department of Civil & Environmental Engineering at Hanyang University, where he served as Dean of the College of Engineering Science from 2002 to 2004. Dr. Cho has been the recipient of various honors and awards from the Korean government and professional societies. In the last 35 years he has served as consultant and technical adviser in various public/governmental projects for a number of ministries of the R.O.K and the metro city of Seoul. He has been a Fellow of the National Academy of Engineering of Korea since 1999. He served as President of the Korean Society of Steel Construction from 2005 to 2007, and of the Computational Structural Engineering Institute of Korea from 1998 to 2000. Active as well in publishing, Dr. Cho is the author of 12 technical books and over 400 papers in international and Korean journals and conference proceedings. These technical papers concern the application of probabilistic methods in structural engineering, particularly in bridge engineering including risk/reliability assessment, life cycle civil engineering for design and maintenance of civil infrastructures among others.

Der Kiureghian, Armen – Dr. Der Kiureghian received his B.S. and M.S. in Civil Engineering from the University of Teheran, Iran, and his Ph.D. in Structural Engineering from the University of Illinois at Urbana-Champaign in 1975. After three years at the University of Southern California, he joined the faculty at the University of California at Berkeley. Since July 1999, he has held the Taisei Chair in Civil Engineering at UC Berkeley. Prof. Der Kiureghian's teaching and research interests are in the areas of structural reliability, risk analysis, stochastic structural dynamics and earthquake engineering. He has pioneered methods for safety and reliability assessment of complex structures and for stochastic dynamic analysis of buildings, bridges and critical equipment. He has more than 300 publications, including more than 100 in archival journals and as book chapters. He is a member of several professional societies including the ASCE, EERI, SSA, IASSAR, and CERRA in which he served as a past President. Currently, he serves on the editorial boards of several journals, including *Probabilistic Engineering Mechanics*, *Structural Safety*, *Earthquake Engineering & Structural Dynamics*. He is the recipient of several awards, including the ASCE Walter L. Huber Research Prize (1988), the CERRA Award (1999), the Movses Khorenatsi Medal from the Government of Armenia for his role in founding the AUA, the ASCE Alfred M. Freudenthal Medal (2006), and the ASCE Thomas A. Middlebrooks Award (2006). In 2004 he was selected as a Distinguished Alumnus of the Faculty of Engineering of Tehran University, Iran, and in 2006 as a Distinguished Alumnus of the Department of Civil and Environmental Engineering of the University of Illinois at Urbana-Champaign. He is an elected foreign member of the National Academy of Engineering and the National Academy of Sciences of Armenia.

Ellingwood, Bruce R. – Dr. Ellingwood is currently the holder of the Raymond Allen Jones Chair in Civil Engineering, Georgia Institute of Technology. Before joining Georgia Tech in 2000, he was Professor of Civil Engineering at Johns Hopkins University. He received his Ph.D. in Civil Engineering from the University of Illinois at UrbanaChampaign in 1972. His areas of research include applications of methods of probability and statistics to structural engineering; structural reliability analysis; structural load modeling and analysis of combinations of loads; development of safety and serviceability criteria for structural design; random vibration; stochastic mechanics; abnormal loads and progressive collapse; response of structures exposed to fires; probabilistic risk assessment of engineered facilities; and performance-based engineering. He has received numerous awards including the N. M. Newmark Medal, Walter P. Moore, Jr. Award, Norman Medal, Moisseiff Award, State of the Art of Civil Engineering Award, Walter L. Huber Research Prize from ASCE, fellow of the ASCE, Lifetime Achievement Award from the AISC, and elected to the National Academy of Engineering in 2001. He is Editor of the Journal of Structural Safety and serves on editorial board of several other journals.

Fisher, John W. – Dr. Fisher was Professor of Civil Engineering at Lehigh University since 1969, and became Professor Emeritus in 2002. He was the founding Director of the NSF Engineering Research Center on Advanced Technology for Large Structural Systems (ATLSS) since its establishment in May 1986 until September 1999. Dr. Fisher is a graduate of Washington University, St. Louis, Missouri, with MSCE and PhD. degrees from Lehigh University. He is a specialist in structural connections, fatigue and fracture of riveted, bolted and welded structures and composite steel and concrete members. He was elected to the National Academy of Engineering, a Corresponding member of the Swiss Academy of Engineering Sciences, and an Honorary Member of ASCE. Other honors include the Engineering Societies John Fritz Medal, the ASCE Ernest E. Howard award, the Franklin Institute Frank B. Brown Medal, the TRB Roy W. Crum Distinguished Service Award and Distinguished Lecturer, the IABSE International Award of Merit in Structural Engineering, the AISC T.R. Higgins Award, the Educator Achievement award and the Geerhard Haaijer award for Excellence in Education. He was ENR's Construction Man of the Year in 1986 and identified in 1999 as one of the top people contributing to the construction industry since 1874. In 2007 he received the Outstanding Projects and Leaders (OPAL) Lifetime Achievement Award in Education from ASCE.

Frangopol, Dan M. – Dr. Frangopol is the first holder of the Fazlur R. Khan Endowed Chair of Structural Engineering and Architecture at Lehigh University, in Bethlehem, Pennsylvania. Before joining Lehigh University in 2006, he was Professor of Civil Engineering at the University of Colorado at Boulder. In 1976, he received his doctorate in Applied Sciences from the University of Liège, Belgium, and holds two honorary doctorates (Doctor Honoris Causa) from the Technical University of Civil Engineering in Bucharest, Romania, and the University of Liège, Belgium. He is a Fellow of ASCE, ACI, and IABSE, Honorary Member of the Romanian Academy of Technical Sciences, and Honorary Member of the Portuguese Association for Bridge Maintenance and Safety. For ASCE he has chaired the Executive and Awards Committees of the Structural Engineering Institute, Technical Activities Division, the Technical

Administrative Committee on Structural Safety and Reliability, and the Technical Committees on Safety of Buildings, Safety of Bridges, and Optimal Structural Design; he is the founder and current chair of the ASCE-SEI Technical Council on Life-Cycle Performance, Safety, Reliability and Risk of Structural Systems. His main areas of expertise are probabilistic methods and structural reliability, lifetime probabilistic optimization of structures and infrastructures, bridge engineering, and life-cycle analysis, design, maintenance, and management of structures and infrastructures. His novel definition of redundancy of structural systems served for the development of the first (1994) AASHTO-LRFD Bridge Design Specifications. Prof. Frangopol is the Founding President of the International Association for Bridge Maintenance and Safety (IABMAS) and of the International Association for Life Cycle Civil Engineering (IALCCE). He is the recipient of several awards including the N.M. Newmark Medal, the T.Y. Lin Medal, the J.J.R. Croes Medal, the E. Howard Award, and the Moisseiff Award. He is the Founder and Editor-in-Chief of the Journal of Structure and Infrastructure Engineering.

Furuta, Hitoshi – Dr. Furuta is a Professor in the Department of Informatics at Kansai University, Osaka, Japan. He received his doctorate in Engineering from Kyoto University, Japan. Before joining Kansai University, he worked for eighteen years in the Department of Civil Engineering at Kyoto University. He was a Visiting Assistant Professor at Purdue University, a Visiting Scholar at Princeton University, and a Visiting Professor at the University of Colorado at Boulder. He is currently the Vice Dean of Faculty of Informatics, Kansai University. His main areas of expertise are in structural reliability, structural optimization, life-cycle cost analysis, design of bridge structures, and applications of soft computing including artificial intelligence, fuzzy logic, neural network, chaos theory, and genetic algorithm. Prof. Furuta was the vice president of IFSA (International Fuzzy Systems Association) and chairman of IFIP (International Federation for Information Processing) Technical Committee 7.5. Currently, he is Chair of the Committee of Structural Engineering, and the Task Committee on Load Specification of the Japan Society of Civil Engineers, chair of the Committee of Asset management in Japanese Construction Consulting Association, and the chair of the committee on Data Exchange Format for Highway Construction in Japan Highway Corporation. He is the author or co-author of over 200 refereed publications. He is also the co-editor of three books published by ASCE and Balkema and an editorial board member of several international journals. Additionally, he has chaired and organized several international structural engineering conferences. He is currently President of IASSAR (International Association of Structural Safety and Reliability), a member of Working Commission 1, IABSE, and Vice President of IALCCE (International Association for Life-Cycle Civil Engineering).

Gilbert, Robert – Dr. Gilbert is the Hudson Matlock Professor in Civil, Architectural and Environmental Engineering at The University of Texas at Austin. He joined the faculty in 1993. Prior to that, he earned B.S. (1987), M.S. (1988) and Ph.D. (1993) degrees in civil engineering from the University of Illinois at Urbana-Champaign. He also practiced with Golder Associates Inc. as a geotechnical engineer from 1988 to 1993. His expertise is the assessment, evaluation and management of risk in civil engineering. Applications include building foundations, slopes, pipelines, dams and levees,

landfills, and groundwater and soil remediation systems. He teaches undergraduate, graduate and professional level courses, and conducts research supported by federal and state agencies and private industry. Recent activities include analyzing the performance of offshore platforms and pipelines in hurricanes; managing earthquake and flooding risks for the Sacramento-San Joaquin Delta in California; and performing a forensic analysis of the New Orleans levee failures.

Haldar, Achintya – Dr. Haldar is currently Professor of Civil Engineering & Engineering Mechanics and da Vinci Fellow at the University of Arizona. He received his MS and PhD degrees from the University of Illinois @Urbana-Champaign. After graduation, he worked for the Nuclear Power Division of Bechtel Power Corporation in Los Angeles. He also taught at Illinois Institute of Technology and Georgia Institute of Technology. He has given short courses at the Hong Kong University of Science & Technology and at the Technical University of Ostrava in the Czech Republic. His main research is on risk and reliability applied to many branches of engineering. So far, he has published over 350 technical articles. Dr. Haldar received many awards for his research, including the NSF Presidential Young Investigator Award and the ASCE's Huber Civil Engineering Research prize. He received an Honorable Diploma from the Czech Society for Mechanics, the Distinguished Alumnus Award from the Civil and Environmental Engineering Alumni Association, University of Illinois. He received the Hind Rattan (Jewel of India) Award. The Arizona Society of ASCE awarded him the John C. Park Outstanding Civil Engineer Award, and was named Graduate Advisor of the Year from the University of Arizona. Dr. Haldar received numerous recognitions for his exceptional teaching.: the Outstanding Faculty Member Award, the Professor of the Year Award and Award for Excellence at the Student Interface several times. He received the Burlington North Foundation Faculty Achievement Award from the University of Arizona. He is a registered professional engineer in several States. An ASCE fellow, Dr. Haldar provides services to the profession at all levels and was recognized several times for his activities.

Holtz, Robert D. – Dr. Holtz is Professor of Civil Engineering at the University of Washington in Seattle. He has degrees in civil engineering from Minnesota and Northwestern, and participated in the Special Program in Soil Mechanics at Harvard under Professor A. Casagrande. His research and publications are mostly on geosynthetics, foundations, soil improvement, soil properties, and instrumentation. He is a Fellow, Life, and Distinguished Member of ASCE, was President of the Geo-Institute of ASCE in 2000–01, and currently serves as International Secretary for the Geo-Institute Board of Governors. Prof. Holtz is a registered engineer in California and Indiana, and throughout his academic career, he has had an active consulting practice. His projects have involved various aspects of geosynthetics and soil reinforcing, foundations, soil improvement, slope stability and landslides, investigation of failures, and acting as an expert witness.

Lee, Seng-Lip – Professor Lee obtained his Ph.D. from the University of California, Berkeley, in 1953. After working in Kaiser Engineers and Bechtel Corporation in California, he joined the Department of Civil Engineering at Northwestern University, Evanston, Illinois, in 1955 as an assistant professor and attained full professorship in

1960. He then accepted the position of professor and chairman of Division of Structural Engineering and Mechanics at the Asian Institute of Technology in Bangkok, Thailand. In 1975, he joined the University of Singapore as Professor and Head of the Department of Civil Engineering. He is currently Emeritus Professor in Civil Engineering at the National University of Singapore. His recent research areas cover structural engineering, geotechnical engineering and construction technology. He has published over 500 papers and has delivered many keynote addresses in many countries. Professor Lee has served as a structural and/or geotechnical engineering consultant in more than 100 projects in several countries involving design and construction of infrastructures such as high rise buildings, high-tech buildings, subway vibration, land reclamation, repair of buildings and bridges. He has served as a director of several companies and is registered as a Professional Engineer in several countries and states in the U.S. He is co-holder of a patent on Fibredrain for consolidation of clayey soils. Professor Lee has received numerous honors and awards including the A.P. Greensfelder Construction Prize (1982) of the ASCE, Maurice P. van Buren Structural Engineering Award (1989) of the ACI, International Federation of Asian and Pacific Contractors Association and Philippines Contractors' Association Foundation Research awards for Works on Deep Basements (1979) and Soil Stabilization (1982), Public Administration Medal (Silver) Singapore (1985), Singapore National Science and Technology Award (1987), First Institution of Engineers Singapore Innovators' Awards (1990) and the University of California, Berkeley Distinguished Engineering Alumnus Award (1991).

Liu, Liang Y. – Dr. Liu is currently Associate Professor in the Department of Civil and Environmental Engineering of the University of Illinois at Urbana-Champaign. He holds a B.S. degree in Civil Engineering from National Chiao-Tung University, Taiwan and M.S. and Ph.D. degrees in Construction Engineering and Management from the University of Michigan, Ann Arbor. Prof. Liu teaches graduate and undergraduate courses in Construction Productivity, Cost Estimating, Construction Management Information Systems, and Construction Case Studies. His research areas include project controls, productivity analysis and improvement, cost engineering, information technologies, and facility life-cycle analysis. He has been recognized several times for excellence in research, teaching and advising. In 2001, he was named the W. E. O'Neil Faculty Research Scholar. In 2003, he won the Teaching Excellence Award of the University of Illinois College of Engineering. He also received the Engineering Council Award for Excellence in Advising in 2008.

Martinez, Julio C. – Dr. Martinez is currently Associate Professor of Civil Engineering at Purdue University. He received his PhD degree from the University of Michigan in 1996, and prior to joining Purdue spent 10 years in the faculty at Virginia Tech. Dr. Martinez' industry experience includes 4.5 years in charge of a utility construction contracting business and 2 years in construction management consulting with Project Management Associates, LL.C. His research foci includes discrete-event simulation modeling, 3D visualization, virtual reality, decision-making under uncertainty, BIM, and construction industry related information technology in general. His discrete-event construction simulation system, Stroboscope, has been used by academics and practitioners worldwide to model complex construction systems and as a basis for

higher-level modeling tools. The Vitascope and Vitascope++ 3D animation and virtual reality systems developed by him and his students define the state-of-the-art in 3D animation of simulated construction operations and virtual interactive construction environments.

Razaqpur, Ghani – Dr. Razaqpur is the Effective Design Chair Professor and Chair of the Department of Civil Engineering at McMaster University in Hamilton, Ontario, Canada. He is also the Director of the Centre for Effective Design of Structures at McMaster. The Centre was founded in 2001 to perform research and train highly qualified personnel in the areas of earthquake engineering, masonry design and materials, use of new and under-utilized materials in construction, and rehabilitation of structures. He is the immediate past president of the Canadian Society for Civil Engineering (CSCE), a fellow of CSCE, a fellow of the Engineering Institute of Canada and a co-editor of the International Journal of Cement and Concrete Composites. Dr. Razaqpur has been involved in teaching, research and consultancy dealing with a large array of topics, including bridge design, analysis and evaluation, design of structures against blast and advanced finite element modeling of structures. He is also been involved in extensive research on durability of structures and design and retrofit of structures using fiber reinforced polymers (FRP). Dr. Razaqpur has published more than 150 conference and journal papers and has supervised 45 master's and doctoral theses.

Shinozuka, Masanobu – Dr. Shinozuka is a Distinguished Professor and Chair of the Department of Civil and Environmental Engineering at the University of California, Irvine (UCI) and also Norman Sollenberger Professor Emeritus at Princeton University. He received PhD in Civil Engineering from Columbia University in 1960, and taught at Columbia, Princeton, and University of Southern California before joining the UCI in 2001. He is a member of the U.S. National Academy of Engineering and an Honorary Member of ASCE. He served as Director of NSF National Center for Earthquake Engineering Research in Buffalo, N.Y. in 1990–1992. His research interest includes continuum mechanics, theory of stochastic processes, structural reliability, and earthquake engineering. Professor Shinozuka is particularly well known internationally for his "Spectral Representation Method" for simulation of multi-variate and multi- dimentional stochastic fields and waves, and for "Scenario-Based Response Analysis Method", both widely applied for simulation-based evaluation of resilience and sustainability of spatially distributed civil infrastructure systems under natural and manmade hazards. He is the recipient of a large number of awards including the T. von Karman, A.M. Freudenthal, and N.M. Newmark medals from the ASCE.

Tang, Wilson H. – Professor Tang is Chair Professor at the Hong Kong University of Science & Technology (HKUST). He received his BS and MS degrees from MIT and Ph.D. from Stanford University. He has taught at the University of Illinois (Urbana-Champaign) for 27 years before returning to his homeland in 1996 to head the Civil Engineering Department at HKUST. His expertise covers applications of probability methods to the wide area of civil infrastructure engineering/management. He has led the profession in promoting the use of reliability-based methods by chairing the USNRC committee on reliability methods for risk mitigation in geotechnical engineering. His selected awards include ASCE Honorary Member, Guggenheim fellow, ASCE

State-of-Art award, US Offshore Energy Center's Hall of Fame, T.K. Hsieh Award (UK), Fellow of Hong Kong Academy of Engineering Sciences, and UIUC Campus Award for Excellence in Undergraduate Instruction. Professor Tang has published over two hundred papers. His co-authored book *Probability Concepts in Engineering Planning and Design* has been widely adopted by top universities around the world. The book has also been translated into more than 5 different languages. He has taken leadership roles on several major international boards and committees in his field, including also local (Hong Kong) professional societies and appointments to government boards.

Vanmarcke, Erik – Dr. Vanmarcke is Professor of Civil and Environmental Engineering at Princeton University. He was on the faculty of the Massachusetts Institute of Technology until 1985, since receiving his doctorate there in 1970. At MIT, he was the Gilbert W. Winslow Career Development Professor and served as the Director of the Civil Engineering Systems Methodology Group. He held visiting appointments at Harvard University, Technical University of Delft (Netherlands), the Catholic University of Leuven (Belgium), his undergraduate alma mater, and was the Shimizu Corporation Visiting Professor at Stanford University. His principal expertise is in engineering risk assessment, reliability analysis, and applied systems science. He authored *Random Fields: Analysis and Synthesis*, published by the MIT Press (2nd edition forthcoming: World Scientific Publishing), and extended this work to the study of space-time processes and complex systems. He was a member of the Independent Panel appointed by President Carter's Science Advisor to review the safety of federal dams. He chaired the Council on Disaster Risk Management and several committees on safety and reliability of the American Society of Civil Engineers (ASCE). He received the Raymond Reese Research Award and the Walter L. Huber Research Prize, both from ASCE, and the Distinguished Probabilistic Educator Award of the Society for Automotive Engineers (SAE). He held a Senior Scientist Fellowship awarded by the Japanese Society for the Promotion of Science and is a foreign member of the Royal Academy of Arts and Sciences of Belgium. An active consultant to government (including, at present, NASA) and industry, he is also founding editor of the international journal *Structural Safety*.

Wen, Yi-Kwei – Dr. Wen is Emeritus Professor of Civil Engineering at the University of Illinois at Urbana-Champaign. He received his D. Eng. Sc. in 1971 from Columbia University and in the same year joined the faculty of Civil Engineering at the University of Illinois at Urbana-Champaign, promoted to full professor in 1981 and has been Emeritus Professor since 2005. He taught courses on probabilistic concepts, structural dynamics, structural reliability and random vibration. His research interests include structural reliability, random vibration, load combination, design based on reliability and lifecycle cost, and applications to buildings and offshore platforms against earthquakes and hurricanes. He was recipient from the American Society of Civil Engineer (ASCE) the Walter Huber Research Prize in 1986, the Moisseiff Award 1986, the Raymond Reese Research Prize in 2002, and the 1997 Research Prize of the International Association of Structural Safety and Reliability (IASSAR). In 2003 he received the Civil Engineering Risk and Reliability Association (CERRA) Award. He is a member and

active in national and international societies (ASCE, AAM, ANSI, EERI, RILEM, IASSAR, IFIP). He has served as consultant to industry in the areas of tornado risks, seismic risks, hurricane and earthquake loads, reliability and reliability-based design of buildings, offshore platforms, and ships. He has published widely including a book, over 70 journal papers, numerous conference papers and editor of international conference proceedings.

Structures and Infrastructures Series

Book Series Editor: Dan M. Frangopol

ISSN: 1747–7735

Publisher: CRC/Balkema, Taylor & Francis Group

1. Structural Design Optimization Considering Uncertainties
   Editors: Yiannis Tsompanakis, Nikos D. Lagaros & Manolis Papadrakakis
   2008
   ISBN: 978-0-415-45260-1 (Hb)

2. Computational Structural Dynamics and Earthquake Engineering
   Editors: Manolis Papadrakakis, Dimos C. Charmpis,
   Nikos D. Lagaros & Yiannis Tsompanakis
   2008
   ISBN: 978-0-415-45261-8 (Hb)

3. Computational Analysis of Randomness in Structural Mechanics
   Christian Bucher
   2009
   ISBN: 978-0-415-40354-2 (Hb)

4. Frontier Technologies for Infrastructures Engineering
   Editors: Shi-Shuenn Chen & Alfredo H-S. Ang
   2009
   ISBN: 978-0-415-49875-3 (Hb)

Forthcoming:

5. Damage Models and Algorithms for Assessment of Structures
   under Operation Conditions
   Editors: S.S. Law & X.Q. Zhu
   2009
   ISBN: 978-0-415-42195-9 (Hb)

Structures and Infrastructures ISSN 1747-7735
Series Editor: Dan M. Frangopol

Frontier Technologies for Infrastructures Engineering

4

Written by eminent world experts in structural engineering, this work presents 19 key contributions on emerging technologies in infrastructure engineering. Focused on the planning, design, construction and maintenance of civil infrastructures, the authors describe how improvements could be implemented in current technologies and how these could benefit. Moreover, they have intended to unite the structural, geotechnical, and construction management aspects.

Specific issues covered in relation to modern structure and infrastructure engineering comprise:

- Dynamics of buildings
- Infrastructures durability
- Corrosion in reinforced concrete
- Design optimization
- Reliability-based design
- Risk assessment and management
- Life-cycle performance and costs
- Health monitoring for maintenance
- Applications in geotechnical and earthquake engineering
- Project management innovations

This volume will serve to advance the technologies for infrastructures engineering everywhere. It is intended for practicing engineers and researchers working on large civil infrastructure systems, such as roadways, bridges, buildings, power generation and distribution systems, water resources and environmental facilities.

CRC Press
Taylor & Francis Group
an **informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
Schipholweg 107C
2316 XC Leiden, NL
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

ISBN-13: 978-0-415-49875-3
9 780415 498753

an **informa** business